Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations 1. Thesis and Dissertation Collection, all items

1984

A preliminary analysis of course loading prediction errors.

Madsen, Craig C.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/19504

Downloaded from NPS Archive: Calhoun

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

A PRELIMINARY ANALYSIS OF COURSE LOADING PREDICTION ERRORS

by

Craig C. Madsen

September 1984

Thesis Advisor: Robert R. Read

Approved for public release, distribution unlimited

T222974

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A Preliminary Analysis of Course Loading Prediction Errors		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis September 1984
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Craig C. Madsen		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93943		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, CA 93943		12. REPORT DATE September 1984
		13. NUMBER OF PAGES 99
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

scheduling, course loading, course predictions, predictions, curricula, courses, median polish, Wilcoxian Signed Rank text, APL, matrix

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The course loading forecast for the Operations Research Department published by the Naval Postgraduate School Programming Office are investigated for measures of predictive accuracy by course and curricula. Both course and curricula forecasts are categorized as problem causing, reasonably accurate, or undefined due to limited information. The paucity of the data precludes the use of sophisticated statistical techniques to

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

20. (Continued)

identify any permanence in effects thereby limiting results to indications of predictive performance. A management aid is developed consisting of data base and manipulative APL programs for the purposes of updating and displaying current information while identifying possible synergistic interactions between data subcatagories. Appendix A contains the Users Guide for this management aid. These results can be applied to the current forecast to aid the department chairman in creating the faculty teaching schedule.

Approved for public release; distribution unlimited.

A Preliminary Analysis of Course Loading Prediction Errors

by

Craig C. Madsen
Lieutenant, United States Navy
B.S., Iowa State University, 1978

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1984

M 274864
c.1

ABSTRACT

The course loading forecasts for the Operations Research Department published by the Naval Postgraduate School Programming Office are investigated for measures of predictive accuracy by course and curricula. Both course and curricula forecasts are categorized as problem causing, reasonably accurate, or undefined due to limited information. The paucity of the data precludes the use of sophisticated statistical techniques to identify any permanence in effects thereby limiting results to indications of predictive performance. A management aid is developed consisting of a data base and manipulative APL programs for the purposes of updating and displaying current information while identifying possible synergistic interactions between data subcatagories. Appendix A contains the Users Guide for this management aid. These results can be applied to the current forecast to aid the department chairman in creating the faculty teaching schedule.

TABLE OF CONTENTS

I. BACKGROUND 10

A. PROBLEM HISTORY 10

B. PRESENT SYSTEM 10

1. Scheduling / Prediction Problems 11

2. Scheduling / Prediction System Use 14

II. INITIAL STEPS AT ANALYSIS 16

A. DATA ORGANIZATION 16

1. Prediction Data 16

2. Historical Data 22

3. Coding System 23

B. NATURAL DIVISIONS IN CODED DATA 24

1. Summary Data 24

2. Curricula Specific Data 25

III. DETAILED ANALYSIS OF OS PREDICTIVE DATA 26

A. THE DATA 26

B. COURSE/CURRICULUM INTERACTIONS AND TESTS . . . 28

C. OTHER INTERACTIONS AND TESTS 33

IV. CONCLUSIONS 36

A. CONCLUSIONS BY COURSE 36

1. Program Output Utilization and Displays 36

2. Display Based Conclusions by Course . . . 40

B. CONCLUSIONS BY CURRICULA 41

1. Program Output Utilization and Displays 41

2. Display Based Conclusions by Curriculum 44

APPENDIX A: PROGRAMS AND DATA BASE USERS MANUAL 47
  A. ACCESS PROCEDURES 47
  B. RUN PREPARATION INSTRUCTIONS 48
  C. SAMPLE PROGRAM RUNS 50
    1. Problem Statement 50
    2. Definition of Goals 51
    3. Course Related Work 51
    4. Curricula Related Work 62
  D. DATA COLLECTION AND MAINTENANCE 70
  E. OTHER AVAILABLE PROGRAMS AND DATA 75

APPENDIX B: SCHEDULING CODES 79

APPENDIX C: DETAILED SUMMARY OF RESULTS 84

APPENDIX D: CONDENSED COURSE MATRICES 89

APPENDIX E: ENTRY / TRACK COURSE DESIGNATIONS 93

LIST OF REFERENCES 96

BIBLIOGRAPHY 97

INITIAL DISTRIBUTION LIST 98

LIST OF TABLES

I. Summary table from [Ref. 1] 13

II. APL Variables used in Analysis (see Appendix A for details of use) 28

III. Wilcoxian Signed Rank Test Results 32

IV. Median Polish Test Results 35

V. Course Results 41

VI. Curricula Results 46

VII. Functions Available 76

VIII. Variables Available 78

IX. Condensed Course Matrices 90

LIST OF FIGURES

2.1 Example of Summary Forecasts (as delivered tc Dept. Chairman) 18

2.2 Example of Individual Course Predictions 20

2.3 Partial Listing of Degree to Curriculum Codes (see [Ref. 3] or Appendix B for complete listing) 21

2.4 Sample Attendance Data (names removed) 23

3.1 Example of Array Format (with definitions) APL variable used : OS 27

3.2 Condensed Course Matrix Examples 30

3.3 Example of Entry vs Track Course Organization 31

4.1 Sample Output from COURSEP Program 37

4.2 Sample Output from COURSE Program Options : Curricula Grouping with Error Percentages . . . 39

4.3 Sample Output from COURSE Program Options : Quarter Grouping with Error Percentages 40

4.4 Sample Output from CURRICP Program 42

4.5 Sample Output from CURRIC Program Options : Quarter Grouping with Error Percentages 43

4.6 Sample Output from CURRIC Program Options : Course Grouping with Error Percentages 45

A.1 Input Syntax : COURSEP 53

A.2 Output from Input Sequence shown in Figure A.1 : COURSEP Program 54

A.3 'COURSEP' Program 55

A.4 Input Syntax : COURSE 56

A.5 Output from Input Sequence shown in Figure A.4 : COURSE Program 57

A.6 'COURSE' Program 58
A.7 Input Syntax : COURSE 60
A.8 Output from Input Sequence shown in Figure A.7 : COURSE Program 61
A.9 Input Syntax : CURRICP 63
A.10 Output from Input Sequence shown in Figure A.9 : CURRICP Program 64
A.11 'CURRICP' Program 65
A.12 Input Syntax : CURRIC 66
A.13 Output from Input Sequence shown in Figure A.12 : CURRIC Program 67
A.14 'CURRIC' Program 68
A.15 Individual Course Prediction Format 72
A.16 Sample Grade Sheet (names removed) 73
A.17 CMS File Format 74
A.18 Wilcoxian Signed Rank Test Program 77

I. BACKGROUND

A. PROBLEM HISTORY

The Naval Postgraduate School uses a complex prediction system to forecast expected student demand for courses over the upcoming year. This forecast is computed using various inputs from several sources during the summer quarter (July to September) and covers the following academic year - fall through summer quarters inclusive. The tabled output of course loading predictions so computed is used by the various Department Chairmen to schedule faculty teaching responsibilities and research time for that year. As with any forecasting system, this set of course loading predicitons is not exact. Various errors are introduced either from the simplification of the process required to get the model to a managable size or from erroneous input. The Operations Research Department Chairman noticed an inaccuracy problem when a class with a predicted size of approximately 30 was scheduled and approximately 75 students showed up, causing a disruption of faculty teaching plans. Believing that a problem existed, he posted a request to research the scope of the inaccuracies as pertaining to the courses taught by the Operations Research Department. The study's goal would be to determine if a significant problem existed and, if so, to provide either a series of adjustments to be made to the present predictions or a better prediction method.

B. PRESENT SYSTEM

Until 1972, the scheduling and course prediction system was run by hand using "corporate memory"; the historical demand for courses as recalled by the curriculum representatives involved. These recollections were used to prepare a proposed schedule that the students and curricular offices used and modified as required. By 1972, this system had just about reached the limits of its capabilities. At that time, the scheduling problem was passed on to a newly hired NPS computer program designer[1] who tailored a predictive computer model to the school's needs. That model, with modifications, is still in use today.

1. Scheduling / Prediction Problems

The problem faced in 1971 is the same today - How can we accurately predict three items:

- The courses required to match student demand
- When those courses should be taught
- How many students will attend

Two types of student inputs are available; those who are not yet onboard (with or without orders), and those presently attending the school in any given capacity. Both types of students are designated as attending a given curriculum, and from that designation a standard program of courses is assigned. This assignment leads to the course scheduling output. Since the two input groups are significantly different with respect to treatment within the model itself, they will be separately discussed.

a. Students not yet onboard

The Office of the Chief of Naval Operations publishes a yearly quota plan, [Ref. 1], which directs the Commander, Naval Military Personnel Command to provide the officer personnel to fill the allocated graduate education billets. The fiscal year 1985 quota summary is shown in table I . These quotas are further broken down to show, by professional designator (i.e. Surface Warfare, NFO, Cryptology . . .), what curricula the officers will attend. Historically, however, these quotas have been guidelines and not strictly followed. Additionally, Naval Military Personnel Command (NMPC) gives only a rough idea of in which quarter an officer will arrive during a given year. This in turn drives the input problem experienced at NPS. Through liason with NMPC, the school identifies as specifically as possible just who has orders to attend NPS either in hand or being written. Quota shortfalls tend to occur in the "technical" curricula due to a lack of available, qualified personnel. Conversly, surpluses occur in the "non-technical" curricula. As with any military organization, operational commitments such as maintaining ships in the Eastern Mediterranian Sea or Indian Ocean adversly affect the arrival of Surface Warfare Officers (who consitute the largest group in the NPS student population).

---

[1] Mike Troian, Special Assistant for Programs; presently working under Code 013: J. Barron, Capt. USN (ret)

The school folds all of this quota information and experience into a "best guess" of the number of students arriving per curriculum per quarter.[2] As this process is being completed, the curricular officers designate the paths these new arrivals will follow. This is done either by service (i.e. USN students take the following courses: . . .) or by past experience (i.e. 28% of incoming students need refresher math followed by . . .). The resulting linking of numbers of students and assigned courses provides one type of input to the scheduling model.

---

[2] Additionally, the programming office contacts individual service and foreign representatives to further refine the expected inputs (since [Ref. 1] pertains only to USN officers). These organizations provide less reliable data mainly because the fiscal year for other US services starts in January. The planning for their personnel transfers is therefore not complete until October or November.

TABLE I

Summary table from [Ref. 1]

FISCAL YEAR 1985

REGULAR GRADUATE EDUCATION QUOTA SUMMARY

	Management and Int'l Affairs *************	Engineering and Science ***********	Total *****
Unrestricted Line	93	334	427
Restricted Line	21	109	130
Staff Corps	108	55	163
Doctoral Studies	*	*	5
	***** 222	***** 498	***** 725

NPS, Monterey	555
Civilian Universities	144
AFIT/JAG School/DIC	21
Doctoral Studies	5
	***** 725

* separate plan - published semiannually

b. Students presently attending NPS

The scheduling process for those students already onboard is somewhat simpler. Since they are already designated as attending a given curriculum over a known period of time, the only area in which significant problems are found is that of elective scheduling. Since the prediction is completed a full year in advance, out-quarter electives may not have been correctly selected by the student. His interests or his required course load may change, thereby changing any electives he has selected. Additionally, the course demand may be such that the course is withdrawn from the list of available courses. In all cases, however, course selection input covering the upcoming year is made by the student through his curricular officer to the scheduling office.

2. Scheduling / Prediction System Use

Once elective selections have been made by the onboard students and the numbers of new inputs per curriculum are entered, the course loading prediction model applies standard course matrixes to generate the overall student course demand. The scheduling system is demand based, meaning this derived demand drives the faculty hiring and scheduling process. The budget process also is based on the course demand. Teaching funds are allocated for all instructors teaching course sections with five or more students. The individual department chairmen receive a listing of student demands by course and quarter from which they create a proposed schedule of courses to be offered. This schedule is further adjusted through the allocation of teaching funds in a process consisting of an approximately month long dialog between the department chairmen, division deans, and administrative representatives from the programming office. The goal of this process is to create an acceptable schedule of courses to be offered balanced between annual operating budget funds, student course demands and faculty research requirements. Through this scheduling process, some courses are dropped and others added, creating some student movement among the courses offered. The goal of the course loading prediction model is to minimize the course changes caused by this scheduling.[3]

---

[3] Mike Troian, the Special Assistant for Programs working from Ingersol Room 375, has routinely conducted accuracy reviews of the course loading prediction model. For FY1983, this review has shown a 90% accuracy in predicting actual courses taught with a less than 1% deviation from total number of courses predicted (his figures). As an indicator of faculty required and expected funding costs, the model is reportedly excellent.

II. INITIAL STEPS AT ANALYSIS

Before any fact-based discussion of "better" predictions could be started, the problem facing the Department Chairman had to be defined. A single historical observation of a severe prediction error that caused problems is not significant considering the number of predictions and observations made. Historical data had to be checked to determine if the problem of inaccurate predictions actually existed or was merely assumed based on a seldom occurring event. Since predictions were made by course, the obvious choice for comparison was the historical course data available from the registrar's office. Not until these steps were completed could a search for a better method of prediction be started.

A. DATA ORGANIZATION

One of the initial steps of the analysis was an organization of all the usable data into a quantifiable format. Useful information available consisted of the following items:

- course number
- quarter taught
- number of students predicted
- number of students attending

The eventual goal of the analysis became that of finding such permanencies in the data (based on these items) as to provide an accurate prediction adjustment.

1. Prediction Data

Prediction data was available both for individual curricula inputs into a given course and for the total expected student demand for those courses. This data was presented in two formats: summary data and individual course data. The summary data (figure 2.1) existed for academic years 1982, 1983 and the present. This data provided a given years display of the total expected student loading indexed by course and quarter. It included a recommended number of "sections taught" and the corresponding faculty hours required. As seen in the figure, the predictions are made for a given academic year, in this case 84 (September 1983 to August 1984). The first column contains the alpha-numeric course identifier as listed in [Ref. 2]. The 'STU' columns refer to the number of students predicted to attend, the 'SEC' column refers to the number of sections required to accomodate those students (maximum section size is limited to 30 attendees), and the 'FAC-HRS' columns refer to the number of faculty teaching hours needed to teach those sections.

```
SUMMARY REPORT FOR OPERATIONS RESEARCH DEPARTMENT              1 AUGUST 1983
ACADEMIC YEAR 84                                               FAL83 TRU SUMR84

                QUARTER 1         QUARTER 2         QUARTER 3         QUARTER 4
              STU SEC FAC-HRS   STU SEC FAC-HRS   STU SEC FAC-HRS   STU SEC FAC-HRS
OA 2200  4 1    0   0   0        43   2  10         0   0   0        34   2  10
OA 2600  4 0   33   2   8         0   0   0        25   1   4         0   0   0
OA 2651  4 0   12   1   4         0   0   0        10   1   4         0   0   0
OA 3101  4 1   46   2  10         0   0   0        35   2  10         0   0   0
OA 3102  4 1    0   0   0        47   2  10         0   0   0        36   2  10
OA 3103  4 1   32   2  10         0   0   0        47   2  10         0   0   0
OA 3104  3 1    0   0   0        32   2   8         0   0   0        45   2   8
OA 3201  4 0   33   2   8        45   2   8         2   1   4        35   2   8
   .
   .
   .
OS 3404  3 0    0   0   0        61   3   9         0   0   0         0   0   0
OS 3601  4 0    0   0   0        26   1   4         0   0   0        14   1   4
OS 3602  4 1   12   1   5         0   0   0        14   1   5         0   0   0
OS 3603  3 1   17   1   4         0   0   0         0   0   0        45   2   8
OS 3604  4 0   27   1   4        44   2   8        20   1   4         1   1   4
OS 3702  4 0    4   1   4        12   1   4         5   1   4        14   1   4
OS 4301  4 0    0   0   0         0   0   0         0   0   0         1   1   4
```

Figure 2.1 Example of Summary Forecasts
(as delivered to Dept. Chairman).

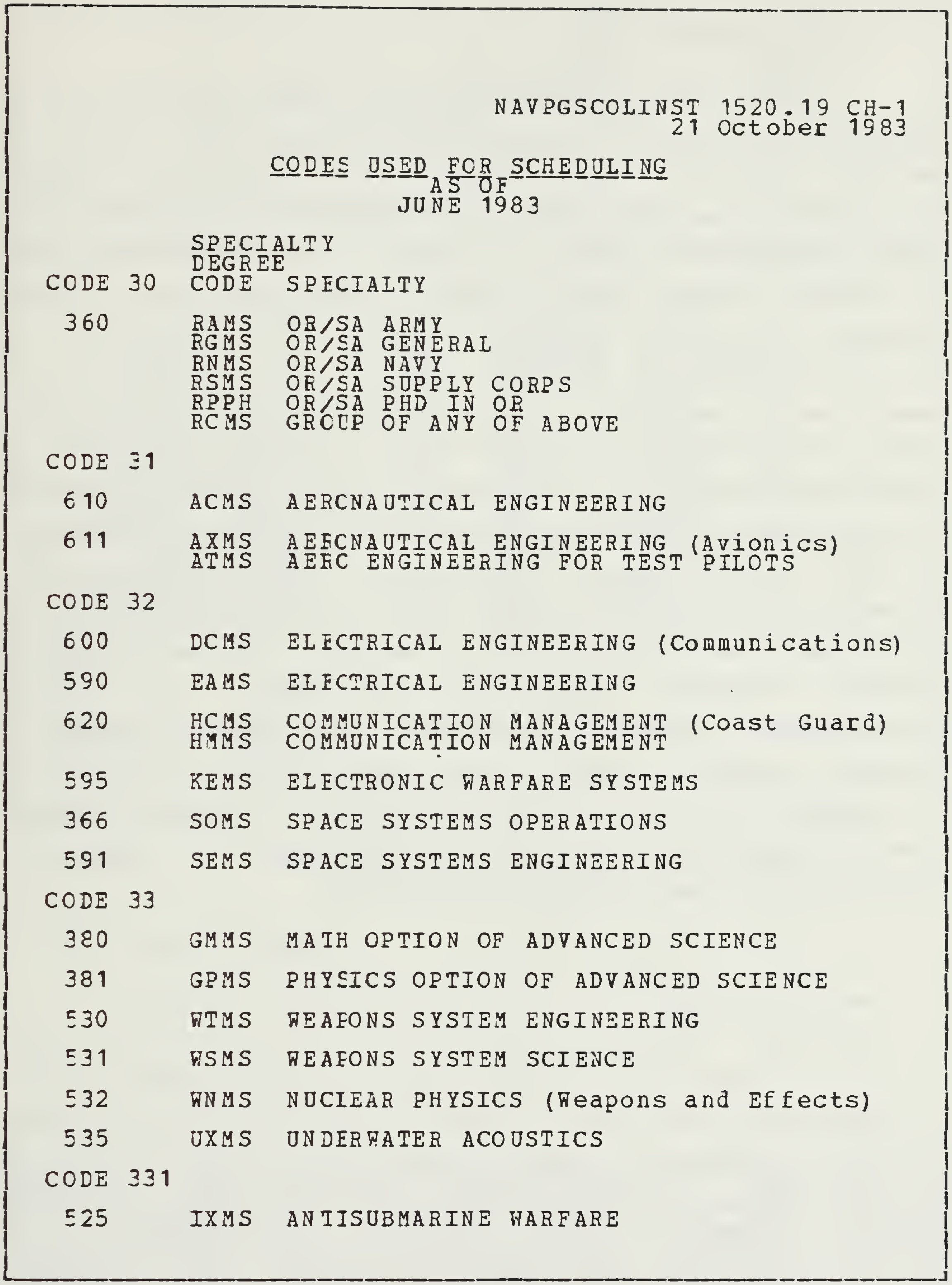
For 1983 and 1984, individual course predictions were available (figure 2.2). These course predictions listed the expected student loading for a given course during a given quarter coded by degree awarded. This degree coding is found in the "S/D/T" (specialty/degree/time) column of the data. The first two letters of this column apply to the curriculum specialty, the next two to the degree (MS refers to Master of Science). This coding is easily transferrable to curriculum number using Naval Postgraduate School instruction 1520.19 (figure 2.3). The numbers that occasionally appear following those four letters indicate when curricular officers have not provided an updated list of whom they expect for that course at that time in that specialty. The numbers mean that the prediction listed came from the "standard files" of the programming office; a collection of best guesses of the expected number of attendees loosely based on historical patterns.[4] These individual course predictions, however, make no mention of segment assignments.

---

[4] Other columns in the individual course prediction data are generally self-explanatory with the exception of these minor items: The AV column referred to the number of aviators (Pilots, NFOs, etc.) that attend and is an unused relic of some past program. The letters sometimes preceding the section column are now meaningless. The section column code itself is assigned by the curricular offices to groups of incoming students.

```
ACADEMIC DEPARTMENT CHAIRMAN REPORT FOR OPERATIONS RESEARCH DEPT  1 AUGUST 1983
ACADEMIC YEAR 84     ACADEMIC QUARTER 1                            FAL83 TRU SUMR84

COURSE LECT/LAB SECTION STUDENTS AV  S/D/T  SEGEMENT INSTRUCTOR  ROOM  TEXTS
*********************************************************************************
OA 2600   4    0    RC3301     5     0   RCMS
                    RG3301     4     0   RGMS
                    RG4101     3     0   RGMS
                  C RG4101     5     1   RGMS01
                    RN3301     2     0   RNMS
                  C RN4101     5     1   RNMS01
                    RN4102     7     0   RNMS
                  C RS4101     2     1   RSMS01

  TOTAL NO. STUDENTS ENROLLED IS 33

OA 2651   4    0 R  RA4101     5     1   RAMS01
                    RA4101     6     0   RAMS
                    RC3302     1     0   RCMS

  TOTAL NO. STUDENTS ENROLLED IS 12


    .
    .
    .


OS 3602   4    1    IX2102     2     0   IXMS
                    IX2103     2     0   IXMS
                    IX2104     2     0   IXMS
                    IX2301     1     0   IXMS
                    IX2302     1     0   IXMS
                    IX2303     1     0   IXMS
                    XT2201     1     0   XTMS
                    XT2303     1     0   XTMS
                    XT3101     1     0   XTMS

  TOTAL NO. STUDENTS ENROLLED IS 12
```

Figure 2.2 Example of Individual Course Predictions



NAVPGSCOLINST 1520.19 CH-1
21 October 1983

CODES USED FOR SCHEDULING
AS OF
JUNE 1983

CODE 30	SPECIALTY DEGREE CODE	SPECIALTY
360	RAMS	OR/SA ARMY
	RGMS	OR/SA GENERAL
	RNMS	OR/SA NAVY
	RSMS	OR/SA SUPPLY CORPS
	RPPH	OR/SA PHD IN OR
	RCMS	GROUP OF ANY OF ABOVE
CODE 31		
610	ACMS	AERONAUTICAL ENGINEERING
611	AXMS	AERONAUTICAL ENGINEERING (Avionics)
	ATMS	AERO ENGINEERING FOR TEST PILOTS
CODE 32		
600	DCMS	ELECTRICAL ENGINEERING (Communications)
590	EAMS	ELECTRICAL ENGINEERING
620	HCMS	COMMUNICATION MANAGEMENT (Coast Guard)
	HMMS	COMMUNICATION MANAGEMENT
595	KEMS	ELECTRONIC WARFARE SYSTEMS
366	SOMS	SPACE SYSTEMS OPERATIONS
591	SEMS	SPACE SYSTEMS ENGINEERING
CODE 33		
380	GMMS	MATH OPTION OF ADVANCED SCIENCE
381	GPMS	PHYSICS OPTION OF ADVANCED SCIENCE
530	WTMS	WEAPONS SYSTEM ENGINEERING
531	WSMS	WEAPONS SYSTEM SCIENCE
532	WNMS	NUCLEAR PHYSICS (Weapons and Effects)
535	UXMS	UNDERWATER ACOUSTICS
CODE 331		
525	IXMS	ANTISUBMARINE WARFARE

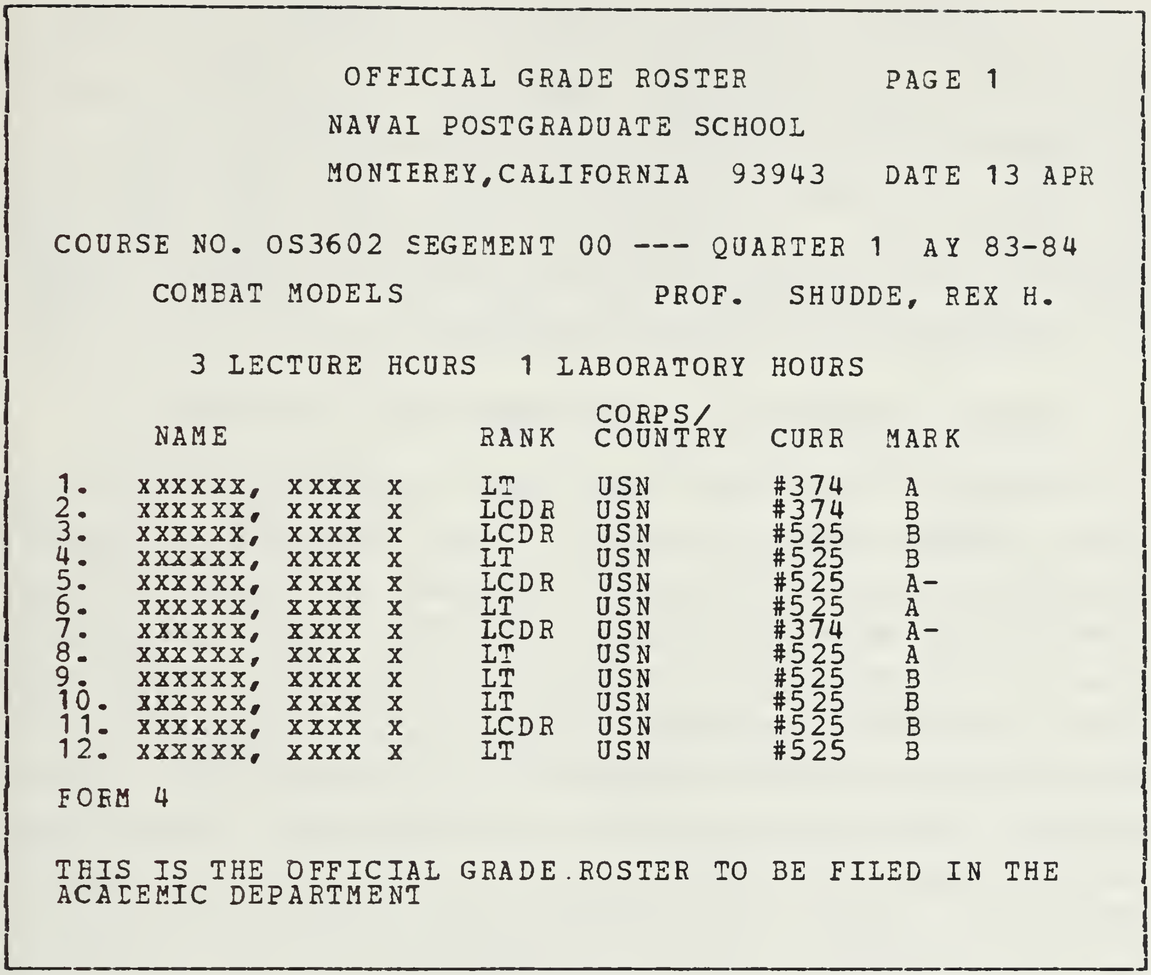
Figure 2.3 Partial Listing of Degree to Curriculum Codes (see [Ref. 3] or Appendix B for complete listing).

2. Historical Data

Historical attendance data was made available by the registrar's office. This data, filed by quarter[5] listed the courses taught in alpha-numeric order displaying student names, curricula, and grades received (see figure 2.4). The numbers recorded reflect the number of students completing the course vice the number seated on the first day of class. Segements are listed separatly. Use of this historical data implies a number of actions. First, the predictions are demand based and would probably correlate better with the number of students registered for the course than with the number completing. A strong positive correlation is assumed between number of students registered and number completing a given course. Second, grade and name information was discarded thereby burying any measure of correlation between performance and attrition. Third, segment data was collapsed to match individual course prediction data. This obscured relationships between predicted and actual number of sections taught or faculty hours required. Since the goal of the analysis was to adjust the predictions of student demand vice to explain the reasons an adjustment is necessary, these actions were considered acceptable.

---

[5] The registrar's system is rather unique in its coding of this data. 'School Year 84' does *not* refer to fiscal year 84 (normally 1 October 1983 - 30 September 1984). Instead, 'School Year 84' refers to the September 84 to August 85 quarters. As this data was collected, the coding of quarterly data (to be discussed later) was translated back into standard fiscal year notation. Thus, throughout the data base used in the analysis, 83.1 refers to the September 82 - December 82 quarter of the 1982 - 1983 school year, as expected.



```
                 OFFICIAL GRADE ROSTER        PAGE 1
                NAVAL POSTGRADUATE SCHOOL
                MONTEREY,CALIFORNIA  93943   DATE 13 APR

 COURSE NO. OS3602 SEGEMENT 00 --- QUARTER 1  AY 83-84
      COMBAT MODELS                PROF.  SHUDDE, REX H.

        3 LECTURE HOURS  1 LABORATORY HOURS

                                  CORPS/
      NAME                  RANK  COUNTRY   CURR   MARK

 1.  xxxxxx, xxxx x         LT    USN       #374    A
 2.  xxxxxx, xxxx x         LCDR  USN       #374    B
 3.  xxxxxx, xxxx x         LCDR  USN       #525    B
 4.  xxxxxx, xxxx x         LT    USN       #525    B
 5.  xxxxxx, xxxx x         LCDR  USN       #525    A-
 6.  xxxxxx, xxxx x         LT    USN       #525    A
 7.  xxxxxx, xxxx x         LCDR  USN       #374    A-
 8.  xxxxxx, xxxx x         LT    USN       #525    A
 9.  xxxxxx, xxxx x         LT    USN       #525    B
 10. xxxxxx, xxxx x         LT    USN       #525    B
 11. xxxxxx, xxxx x         LCDR  USN       #525    B
 12. xxxxxx, xxxx x         LT    USN       #525    B
 FORM 4

 THIS IS THE OFFICIAL GRADE.ROSTER TO BE FILED IN THE
 ACADEMIC DEPARTMENT
```

Figure 2.4   Sample Attendance Data (names removed)

3. Coding System

Before entering the data into computer files, a coding system was devised to distill the raw data into its usable parts. Excess information such as student names and grades could then be discarded. The data consisted of predictions and actual attendance figures coupled with identifiers. Identifiers required include course number and quarter taught. Available data included both total and individual curricula figures. For simplicity of use, a basic file was created using the following five-column coding scheme:

Quarter Curricula# Course# #attended #predicted

Once these items had been tabulated, an easy subtraction (attending minus predicted) provided an error term. By ordering (by magnititude of error) and displaying the resulting array, a history of many poor predictions was observed. This observation confirmed that a problem existed and showed a need for further analysis.

B. NATURAL DIVISIONS IN CODED DATA

The large volume of coded data created was divisible into several sub-categories. The Operations Research Department teaches two types of courses: OA - taken by operations analysis major students; and OS - taken by students from other curricula. Initially both of these types were taken under consideration for analysis. Later, the OA designated courses were dropped from consideration for two reasons; the problem of not knowing how many students were going to show up was generally solvable within the department; additionally, time constraints limited the scope of the areas to be covered.[6] The OS course data was broken into two prediction areas: summary predictions and curricula specific predictions. The summary predictions were used to show patterns by course; the curricula predictions to show patterns by individual curricular offices.

1. Summary Data

When dealing with summary prediction data error terms, the goal of the analysis was to provide a functional relationship between predictions and attendance by course.

---

[6] OA data collected during the initial stages of problem investigation is available. The format is slightly different from that of the OS course data because the problem found for OA dealt with elective selection. The factor of interest was service affiliation. For further details, see appendix A.

Using these derived relationships (hopefully simple additions or subtractions), the department chairman will be able to take the forecast from the programming office (figure 2.1) and change the figures shown to reflect more accurately the expected student load. These more accurate figures can then be used to schedule faculty teaching assignments, request necessary funding, and accomplish other tasks requiring the number of expected students per course.

2. Curricula Specific Data

When dealing with error terms for individual curricula, the goal of the analysis was to provide a functional relationship between predictions and attendance by curriculum. Using these derived relationships, the department chairman will be able to take the forecast for an individual course (figure 2.2) and change the figures shown to reflect more accurately the expected student load by either applying the relationship or calling the curricular office involved to get better numbers. These historical observations will enable the department chairman both to identify curricula that are inadvertently causing scheduling problems and to show improvements or deterioration in the scheduling process itself. This information can be used both by the programming office and the curricular office involved as well as the OR Department.

III. DETAILED ANALYSIS OF OS PREDICTIVE DATA

At the completion of the initial steps toward analysis previously discussed, a coded series of data arrays were available for use. An example of array formatting is shown in figure 3.1 The figure of interest was the "error" term found in column six. These error terms, representing the difference between the actual number of students attending a course vs the number predicted, were displayed in various groups ordered by magnitude to show the accuracy patterns of the forecast. Areas that were explored for effects included the curriculum that provided the input, the courses themselves, and the quarter for which the courses were predicted. Underlying patterns such as an individuals "path" through various OS courses required for his curriculum were also investigated. The paucity of the data precluded the use of sophisticated statistical techniques thereby limiting results to indications of areas of concern rather than definitive statistical statements.

A. THE DATA

Six element data arrays (as shown in figure 3.1) were created for each years data. These APL variables were combined to form a single compilation of all the available data (see table II). The error terms of this compilation were used for the analysis. These error terms are discrete integer data the size of which are limited by the maximum of the actual or predicted columns terms. Distributional fitting was not done for two reasons; the population should have high internal correlation due to the interrelation between course errors caused by student course matrices, and

```
  QTR   CURRIC   CSE   ACT   PRE   (ACT-PRE)
  ****  ******   ***   ***   ***   *********

  84.3      0    2103    0     1       -1
  84.3    535    2103    0     1       -1
  84.3      0    3001   25     0       25
  84.3    368    3001   25     0       25
  84.3      0    3002   15    18       -3
  84.3    683    3002    0     1       -1
  84.3    825    3002   15    17       -2
   .
   .
   .
  83.1      0    3602    7     7        0
  83.1    374    3602    1     1        0
  83.1    525    3602    6     6        0
  83.1      0    3603   20    21       -1
  83.1    595    3603   20    21       -1
  83.1      0    3604   18    16        2
  83.1    365    3604    1     0        1
   .
   .
   .
```

QTR: Quarter taught. 83.1 refers to the first quarter in the Sept. 82 to Aug. 83 school year.

Curric: Curricula - 0 implys summary data for a given course.

Cse: Course number (all OS courses).

Act: Actual number of students listed on final grade records as completing course.

Pre: Predicted number of students (from forecast).

Act-Pre: Prediction error term.

Figure 3.1 Example of Array Format (with definitions) API variable used : OS.

the sample sizes of those homogeneous populations distilled from the data (such as "all 365 students taking OS 3604 in the second quarter") were prohibitively small. Two types of errors were discovered, predictive errors and accuracy errors. A predictive error was defined as either the forecast predicting no students and students attending or vice versa. Embedded in this error type are those courses which were not even listed but taught and those courses dropped

TABLE II

APL Variables used in Analysis
(see Appendix A for details of use)

VARIABLE ********	SHAPE *******	DESCRIPTION ************
OS82	69x6	AY (Academic year) 81-82 data; includes summary information only (no data for individual curricula)
OS83	255x6	AY 82-83 data; complete listing
OSTRK83	140x8	AY 82-83 non-entry course data *
OS84	228x6	AY 83-84 data; only qtrs 1,2 and 3
OS	552x6	Combination of all of above (without repetition)

* see 'Course/Curriculum Interactions and Tests' section for complete description.

from the schedule in the budgeting process. The same division was found by individual curricula (for example, curriculum 460 (engineering science - preparatory) is never predicted in any individual course forecast). Predictive errors exist when columns five or six of the data array contain a zero. Accuracy errors are other differences between the actual and predicted number of students.

B. COURSE/CURRICULUM INTERACTIONS AND TESTS

Each individual student at the Naval Postgraduate School follows an assigned course matrix. This assignment varies by both the program assigned and individual student needs. However, all matrices list a series of courses that are either required or elecive. The Operations Analysis curriculum, through its OS course offerings, provides a service to other programs by providing basic and advanced pobability and statistics training not available in other curricula. These service courses are generally required by other curricula as a prerequisite to both further training and degree completion. As such, they are taught in a certain sequence. The amount of training required differs with different programs, so various OS "tracks" are required. Sequencing of courses in this manner will cause dependence between the error terms of certain courses. To determine the dependancy relationship within the data, the course matrices for all curricula offered at the school were collected and the OS course sequences extracted. These sequences were combined to form a series of condensed course matrices as applicable to the Operations Research Department (see figure 3.2 - Note: x's refer to students onboard but not taking OS or OA designated courses). From these matrices, patterns emerged. As with any program, students started with basic courses and moved to more advanced topics. Five basic courses were defined as "entry level" courses. One of these courses was required before any advanced work was done. Since the entry level courses were generally taken early in a student's time in his program, and since the entry course was the first time he appeared eligible for "prediction", the entry level courses were expected to have the highest variance in error terms. This proved to be true. As the students spend more time in their programs, attrition and schedule changing should level off. Additionally, since most students spend more than a single year going through their programs, "down-line" predictions will be improved merely by the students moving from the "not yet onboard" to the more stable "attending NPS" prediciton group.

```
                                  QUARTER                             SCHEDULED
CURRIC      1      2      3      4      5      6      7      8      9 ENTRY QTRS
****** ****** ****** ****** ****** ****** ****** ****** ****** ****** **********

 365   OS2103 OS3604 OS3008 OS3603 OS4602                             f
              OS3404

 366   OS2103 OS3604 OS3008 OS3603 OS4602                             f
              OS3404        OA3602

 374   xxxxxx OS2103 OS3604 xxxxxx xxxxxx OS3601 OS3602 OS3603        f,w,sp,su

 525   xxxxxx OS2103 OS3303 OS3601 xxxxxx OS3402 OS4601               f,sp
                     OS3604                      OS3602

 535   xxxxxx xxxxxx xxxxxx OS2103                                    f

 595   xxxxxx OS2103 OS3604 OS3003 OS3603 OS3403 OS4601               f

 367   xxxxxx OS3101 OS3004                                           f,sp

 620   xxxxxx OS3101 OS3005                                           f

 686   xxxxxx OS3101                                                  w,su

 687   OS3101                                                         w,su

 825 A xxxxxx OS3101 OS3002                                           f
              OS3404

     B xxxxxx OS3101 xxxxxx OS3404 OS3002                             sp

 368   OS3001                                                         f,sp

 372   xxxxxx xxxxxx OS3104                                           f,w,sp,su

 373   xxxxxx OS3104                                                  f,w,sp,su

 440   xxxxxx xxxxxx OS3104                                           f,w,sp,su

 441   xxxxxx xxxxxx OS3104                                           f,w,sp,su
```

Figure 3.2 Condensed Course Matrix Examples

Definition of entry level courses led to the definition of course sequence "tracks". A student from a given curricula will take an "entry" course - the first time he is seen by the OA Department - and will then take follow-on courses based on his course matrix. Each curriculum has its own "track" (See figure 3.3). Using these tracks, it should have been possible to derive deterministically the exact course loading for those follow on or track courses given the course matrixes and attendance figures for the entry level courses. Using the FY83 data, this was done. The resulting matrix was set up in the following manner:

ENTRY COURSE	CURRICULUM	TRACK (Follow-on) COURSES, ALL OS
OS2103	365	3604,3404,3008, 3603,4602
	366	3604,3404,3008, 3603,4602
	374	3601,3602,3603, 3604
	525	3303,3604,3601, 3401,4601,3602
	535	none
	595	3604,3003,3603, 3403,4601
OS3101	367	3004
	620	3005
	686	none
	687	none
	825	3404,3002

Figure 3.3 Example of Entry vs Track Course Organization.

Quarter Curric Cse Act Pre Dpre (Act-Pre) (Act-Dpre)

Act : Actual number attending

Pre : Standard predictions

Dpre : Derived predictions

A display of the results showed no observable difference between the regular error term and the derived error term. A Wilcoxian signed rank test[7] was done to test the hypothesis that the error magnitude of the adjusted predictions tends to be smaller than the magnitude of the error from the forecast (see table III).

TABLE III

Wilcoxian Signed Rank Test Results

Sensitive assumption: Differences independent. Assumed large amounts of varied data negated track effects resulting in relatively independent difference terms.

Variables Tested:	CASE A	CASE B
X =	Curric ACT-DPRE	Sum ACT-DPRE
Y =	Curric ACT-PRE	Sum ACT-PRE

H_0 : The X values do not tend to be smaller than the Y values.

H_1 : The X values tend to be smaller than the Y values.

Decision Rule : Reject at the 5% confidence level if TS > 1.65

	CASE A	CASE B
Test Statistic :	.453	-1.455

Conclusion : Cannot reject null hypothesis.

---

[7] Taken from [Ref. 5 p. 280].

The results of the test indicated that the adjusted predictions were no improvement over the forecast. That conclusion indicates that interactions more complex than the following of course matrixes are involved in actual student movement from course to course. The adjusted prediction method was dropped from the analysis as a means of better predicting course loading.

C. OTHER INTERACTIONS AND TESTS

Interactions exist in areas separate from the dependencies created within a class sequence by the course matrix structure. Three main effects were studied. First, curricula may have individual patterns across time. Some groups may consistently have an optimistic input that causes an overestimation of number of students wherever those groups appear. Second, individual courses may show patterns of over- or underestimation. Lastly, a quarterly effect on errors may appear due strictly to the elapsed time between the prediction and the actual class. To provide an initial look at the data to show such patterns, two programs were written.[8] These programs were designed to sort the data into the catagories of interest, specifically curriculum, course, and quarter. The resulting groupings showed that there were patterns in each catagory, but that the number of observations available to base conclusions upon made those patterns statistically insignificant. Through clever data selection, a series of observations was gleaned to test for curriculum and course effects using median polish technique as detailed in [Ref. 4]. The results of that test were inconclusive showing no temporal stability (see table IV). The test for quarterly effects was not done due to the paucity of the data.[9]

---

[8] These programs and others used in the analysis are fully explained and documented in appendix A. That appendix also contains instructions for the maintenance and use of the data base created, and procedures for gaining access to same.

---

[9] Complete data existed only for FY83. This limited sample was evaluated as too unrepresentative to be statistically significant concerning quarterly effects across various years.

TABLE IV

Median Polish Test Results

Two similar matrixes of error data were tested for consistancy of effect. Data format is a/b where multiple observations exist. These observations represent data from different quarters. The APL programs used to provide the listed results were 'MEDPCLISH', 'CODERES' and 'SHOWRES' from public workspace OA3660.

Course	Curriculum 813	814	815	817	819	827	837	847	857	Course Effect A	B
3102	-1	-1	-6/3	2/2	-1	-4/-3	-8/0	-13/0	-3/3	-2	0
3103	-1	0	-6/3	-1/5	-1	-4	-2	0	3	0	0
3006	0	5	1/-3	4/-5	1/-1	1/-3	-3/4	-1/1	5/-4	2	0
Curricula effect											
A	0	2	-3	3	0	-1	-4	-2	4		
B	-1	0	3	2	-1	-3	0	0	3		

	MATRIX A	MATRIX B
Error reduction	From 73 to 40	From 59 to 41 to 38
Typical value	-1	0
Correlation of comparison values with residuals	.123	1
Residual pattern	none	all residuals equal 1

Conclusions : The median polish technique works well when dealing with the 'correct' data (ie. a full, singular entry rectangular matrix). In the rare case where this type of data existed, results greatly differred in cases that should be similar in effect.

IV. CONCLUSIONS

The problem as defined was to aid the Department Chairman in his task of scheduling and budgeting for faculty teaching responsibilities by providing either a better prediction method or adjustments to the present forecasts. From the scheduler's standpoint, jumps across any multiple of thirty students cause new sections to be created or dropped, and courses with a demand of less than five students are dropped. To the scheduler or budget office, these are significant numbers. The Department Chairman, however, cares little about the actual number but much about the accuracy of the prediction. Since predictions are made by course and broken down within individual courses by curriculum, adjustments were found for courses and curricula. Prediction data were the limiting quantities; predictions by curricula exist only as far back as FY83. This limited the scope of the investigation to 1 and 3/4 years. With this limited history, indications of problem or non-problem areas are the only adjustment/accuracy measures possible.

A. CONCLUSIONS BY COURSE

1. Program Output Utilization and Displays

Indications of predictive accuracy by course were obtained using the 'CCURSEP' and 'COURSE' programs detailed in Appendix A. The 'COURSEP' (standing for course patterns) program was used first. This program takes the input matrix, in this case the variable 'OS' containing all of the data so far available, and displays it grouped by course in decreasing order of predictive error magnitude (see figure

```
 COURSE : 3602
    -82.1      0     3602       7       11      -4
    -83.3    374     3602       3        5      -2
    -84.3    374     3602       2        4      -2
    -84.3    525     3602      12       10       2
    -82.3      0     3602      15       14       1
    -83.3      0     3602      15       16      -1
    -83.3    686     3602       1        0       1
    -83.1      0     3602       7        7       0
    -83.1    374     3602       1        1       0
    -83.1    525     3602       6        6       0
    -83.3    525     3602      11       11       0
    -84.1      0     3602      12       12       0
    -84.1    374     3602       3        3       0
    -84.1    525     3602       9        9       0
    -84.3      0     3602      14       14       0
```

Figure 4.1   Sample Output from COURSEP Program

4.1 for a single course example).[10] This output is quickly scanned for 'interesting' courses - those with either significant problems or very small, infrequent errors. Those courses with problems are further checked to determine if there is any pattern to the errors observed. Many negative numbers will indicate an overestimation problem (prediction greater than actual attending), conversely positive numbers indicate underestimation. Quarters may be split indicating large errors for some, small for others. Curricula may be grouped together indicating prediction performance differences. In the example shown (Figure 4.1), error magnitude is small and signs vary, indicating relatively minor deviations from the predicted values. Had these errors been large, the following observations could have been made:

---

[10] In this display (and all others), the quarter is shown as a negative number. This is done strictly as a formatting aid. In APL, a space is added to each column in the output display to allow negative signs. The error terms occasionally had no negatives. When this occurred, the output display deleted the six spaces 'saved' for the sign. This caused an irritating misalignment in the columns of the display. The easy correction was to ensure a negative existed in all cases, thus the quarter coding.

- Third quarter errors tend to be higher than first quarter errors.

- No curricula appears to do any better or worse than any other, though 374 may tend to overestimate in their predictions.

To further detail predictive behavior within a course, the 'COURSE' program is utilized. This program takes the same matrix previously investigated (in this case 'OS') and also displays it grouped by course, but with various catagorical presentations made. Output may be displayed in quarterly or curricula subcategories (through answering program prompts). With the curricula option selected, the output compares patterns of individual curricula (see figure 4.2). Statistics displayed include the mean and variance of the error terms for individuals curricula, and predictive, accuracy and no error percentages also by curricula. To avoid counting data twice, summary data entries (column 2 = 0) are disregarded. From scan of the example (figure 4.2), the following conclusions can be drawn:

- Curriculum 525 provides the majority of the input for OS3602, and their predictions are excellent indicating no problem with the course.

- Curriculum 374 may show a pattern of overestimation.

- Curriculum 686 (with 100% prediction errors) normally will not provide inputs to this course.

```
COURSE : 3602
   -82.1     0     3602      7       11      -4
   -82.3     0     3602     15       14       1
   -83.3     0     3602     15       16      -1
   -83.1     0     3602      7        7       0
   -84.1     0     3602     12       12       0
   -84.3     0     3602     14       14       0

   -83.3   374     3602      3        5      -2
   -84.3   374     3602      2        4      -2
   -83.1   374     3602      1        1       0
   -84.1   374     3602      3        3       0

   -84.3   525     3602     12       10       2
   -83.1   525     3602      6        6       0
   -83.3   525     3602     11       11       0
   -84.1   525     3602      9        9       0

   -83.3   686     3602      1        0       1

  GROUP     MEAN   VARIANCE   N   (OF ERROR TERM)
  374.0    -1.0     1.3      4.0
  525.0      .5     1.0      4.0
  686.0     1.0     1.0      1.0

OVERALL:     -.1    1.6    9.0

  ERROR PERCENTAGES
 GROUP     PRE      ACC     NONE     N
 374.00    .00      .50      .50    4.00
 525.00    .00      .25      .75    4.00
 686.00   1.00      .00      .00    1.00
```

Figure 4.2 Sample Output from COURSE Program Options : Curricula Grouping with Error Percentages.

The curricula option of the 'COURSE' program appears to be the most valuable display available. The information shown is directly transferrable to the prediction data as presented (refer back to figures 2.1 and 2.2).

With the quarter option selected, the output combines all of the data into four catagories depending on the quarter in which the course has been taught (figure 4.3). The same statistics are displayed, this time for each quarter. Again, summary data entries are disregarded to avoid duplication of data. From a scan of this figure, the following conclusions can be drawn:

```
COURSE : 3602
  -82.1      0    3602      7      11      -4
  -83.1      0    3602      7       7       0
  -83.1    374    3602      1       1       0
  -83.1    525    3602      6       6       0
  -84.1      0    3602     12      12       0
  -84.1    374    3602      3       3       0
  -84.1    525    3602      9       9       0

  -83.3    374    3602      3       5      -2
  -84.3    374    3602      2       4      -2
  -84.3    525    3602     12      10       2
  -82.3      0    3602     15      14       1
  -83.3      0    3602     15      16      -1
  -83.3    686    3602      1       0       1
  -83.3    525    3602     11      11       0
  -84.3      0    3602     14      14       0

   GROUP    MEAN  VARIANCE  N  (OF ERROR TERM)
    1.0      .0      .0    4.0
    3.0     -.2     3.2    5.0

OVERALL:     -.1    1.6   9.0

   ERROR PERCENTAGES
  GROUP     PRE     ACC    NONE     N
   1.00     .00     .00    1.00   4.00
   3.00     .20     .60     .20   5.00
```

Figure 4.3 Sample Output from COURSE Program Options : Quarter Grouping with Error Percentages.

- OS3602 has only been taught during the fall and spring quarters.
- Fall quarter predictions tend to be exact.
- Spring quarter predictions are worse, but not to such an extent as to cause problems.

2. Display Based Conclusions by Course

Using the previously discussed displays, the courses were separated into three catagories of interest: those with indications of significant prediction accuracy problems, those with relatively minor problems if any, and those with too little information to draw conclusions (see table V). Predictions made for courses can be checked against this listing for an indication of their expected performance. Should more detailed information be needed, Appendix C contains the results based on summary data from 81-82, complete 82-83 data, and fall, winter and spring quarters from the 83-84 school year.

TABLE V

Course Results

Problem Courses			'Good' Courses		
2101	2102	2103	3002	3004	3006
3001	3008	3101	3303	3402	3601
3102	3103	3104	3602	3603	3604
3105	3404	3702	4601	4701	

Inadequate Data		
3003	3005	3302
3401	3403	4602

B. CONCLUSIONS BY CURRICULA

1. Program Output Utilization and Displays

Indications of predictive accuracy by curricula were obtained using the 'CURRICP' and 'CURRIC' programs detailed in Appendix A. The 'CURRICP' (standing for curricula patterns) program was used first. This program takes the input matrix, 'OS' as before, and displays it grouped by curricula in decreasing order of predictive error magnitude (see figure 4.4 for a single curriculum example). This

```
CURRICULUM : 374
   -83.4   374    2101      0       6      -6
   -83.4   374    2103      2       6      -4
   -83.2   374    2103      1       4      -3
   -83.3   374    3604      1       4      -3
   -83.4   374    3601      2       5      -3
   -83.3   374    3602      3       5      -2
   -83.4   374    3603      4       6      -2
   -84.1   374    3604      2       4      -2
   -84.3   374    3602      2       4      -2
   -83.2   374    3601      4       5      -1
   -83.2   374    3604      0       1      -1
   -84.1   374    2103      0       1      -1
   -84.1   374    3401      0       1      -1
   -84.2   374    2103      2       3      -1
   -84.2   374    3601      3       4      -1
   -84.3   374    3604      2       3      -1
   -83.1   374    3602      1       1       0
   -83.1   374    3604      4       4       0
   -84.1   374    3602      3       3       0
   -84.1   374    3603      1       1       0
```

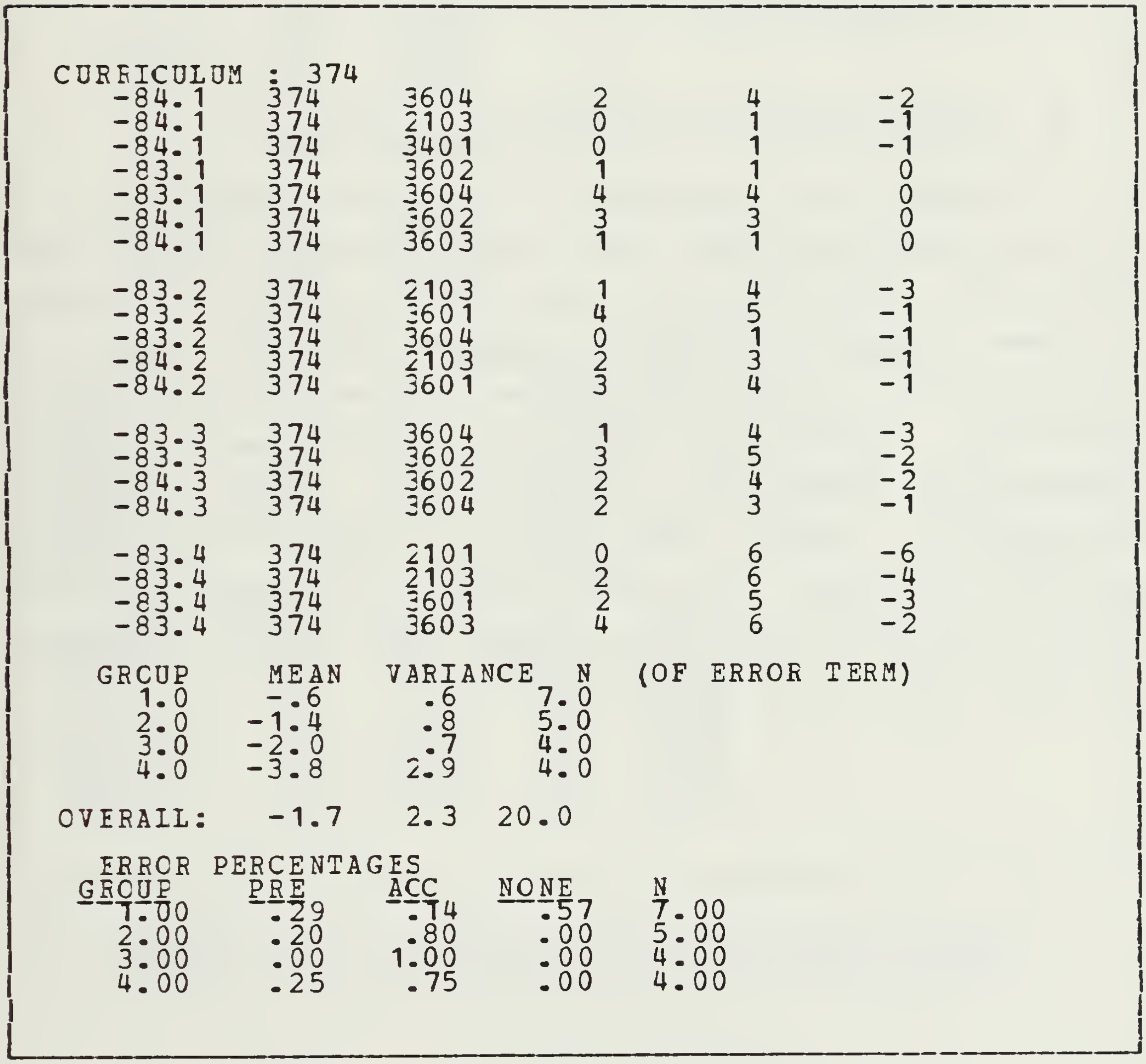
Figure 4.4   Sample Output from CURRICP Program

output is quickly scanned for 'interesting' courses - those with significant problems or small, infrequent errors. Those curricula with significant errors are further checked for discernable patterns such as regularly occuring over- or underestimation, or quarterly trends. In the example shown (figure 4.4), problems appear due to consistent overestimation. However, given the magnitude of the errors, these may be viewed as insignificant. It also appears that fall and winter quarters tend to have fewer or smaller prediction errors.

To more closely investigate patterns shown, the 'CURRIC' program is utilized. This program takes the same input matrix and displays it further grouped by quarterly or course catagories. With the quarter option selected, the output compares predictions across time (see figure 4.5). From the example shown, the following observations can be made:

- Overestimation errors do increase across consecutive quarters.

- Lower level courses tend to have larger magnitude errors than more advanced courses.
- OS2103 and OS3401 might not be taken by curriculum 374 students fall quarter (more data required)



```
CURRICULUM : 374
   -84.1    374     3604       2          4        -2
   -84.1    374     2103       0          1        -1
   -84.1    374     3401       0          1        -1
   -83.1    374     3602       1          1         0
   -83.1    374     3604       4          4         0
   -84.1    374     3602       3          3         0
   -84.1    374     3603       1          1         0

   -83.2    374     2103       1          4        -3
   -83.2    374     3601       4          5        -1
   -83.2    374     3604       0          1        -1
   -84.2    374     2103       2          3        -1
   -84.2    374     3601       3          4        -1

   -83.3    374     3604       1          4        -3
   -83.3    374     3602       3          5        -2
   -84.3    374     3602       2          4        -2
   -84.3    374     3604       2          3        -1

   -83.4    374     2101       0          6        -6
   -83.4    374     2103       2          6        -4
   -83.4    374     3601       2          5        -3
   -83.4    374     3603       4          6        -2

  GROUP      MEAN   VARIANCE   N   (OF ERROR TERM)
    1.0      -.6       .6     7.0
    2.0     -1.4       .8     5.0
    3.0     -2.0       .7     4.0
    4.0     -3.8      2.9     4.0

OVERALL:    -1.7      2.3   20.0

  ERROR PERCENTAGES
 GROUP      PRE      ACC     NONE     N
   1.00     .29      .14      .57    7.00
   2.00     .20      .80      .00    5.00
   3.00     .00     1.00      .00    4.00
   4.00     .25      .75      .00    4.00
```

Figure 4.5 Sample Output from CURRIC Program Options : Quarter Grouping with Error Percentages.

With the course option selected, information can be gleaned on curriculum 374's prediction performance for any given course (see figure 4.6). Again from the example shown, conclusions can be drawn:

- OS2101 and OS3401 are generally not taken by students from curriculum 374
- Larger errors occur in OS2103 and OS3604 (as expected from the course matrix - refer back to figure 3.2)

2. Display Based Conclusions by Curriculum

Using the previously discussed displays, the curricula were separated into three catagories of interest: those with indications of significant prediction accuracy problems, those with relatively minor problems if any, and those with too little information to draw conclusions (see table VI). Predictions made for courses can be checked against this listing for an indication of their expected performance. Should more detailed informtion be needed, Appendix C contains the results based on summary data from 81-82, complete 82-83 data, and fall, winter and spring quarters from the 83-84 school year.

```
CURRICULUM : 374
   -83.4    374     2101       0         6        -6

   -83.4    374     2103       2         6        -4
   -83.2    374     2103       1         4        -3
   -84.1    374     2103       0         1        -1
   -84.2    374     2103       2         3        -1

   -84.1    374     3401       0         1        -1

   -83.4    374     3601       2         5        -3
   -83.2    374     3601       4         5        -1
   -84.2    374     3601       3         4        -1

   -83.3    374     3602       3         5        -2
   -84.3    374     3602       2         4        -2
   -83.1    374     3602       1         1         0
   -84.1    374     3602       3         3         0

   -83.4    374     3603       4         6        -2
   -84.1    374     3603       1         1         0

   -83.3    374     3604       1         4        -3
   -84.1    374     3604       2         4        -2
   -83.2    374     3604       0         1        -1
   -84.3    374     3604       2         3        -1
   -83.1    374     3604       4         4         0

  GROUP      MEAN   VARIANCE   N   (OF ERROR TERM)
 2101.0     -6.0     1.0      1.0
 2103.0     -2.3     2.3      4.0
 3401.0     -1.0     1.0      1.0
 3601.0     -1.7     1.3      3.0
 3602.0     -1.0     1.3      4.0
 3603.0     -1.0     2.0      2.0
 3604.0     -1.4     1.3      5.0

OVERALL:      -1.7    2.3   20.0

  ERROR PERCENTAGES
 GROUP      PRE       ACC      NONE      N
2101.00    1.00       .00       .00     1.00
2103.00     .25       .75       .00     4.00
3401.00    1.00       .00       .00     1.00
3601.00     .00      1.00       .00     3.00
3602.00     .00       .50       .50     4.00
3603.00     .00       .50       .50     2.00
3604.00     .20       .60       .20     5.00
```

Figure 4.6 Sample Output from CURRIC Program Options : Course Grouping with Error Percentages.

TABLE VI

Curricula Results

Problem Curricula

360 368 374 460 530
531 590 817 819 825
827 837 847 857

'Good' Curricula

365 366 367 377 525
595 600 620 813 814
815 999

Inadequate data

372 373 380 440 441
532 535 570 591 610
611 683 686 687

APPENDIX A

PROGRAMS AND DATA BASE USERS MANUAL

This appendix contains the detailed information required to effectively run and maintain the programs and data base used for historical validation of the Naval Postgraduate School course loading forecasts. The purpose of this series of programs is to conveniently display historical course attendance data versus predicted attendance data. These displays can then be translated into simple adjustments to be made to the forecast, providing a more accurate figure to be used in the faculty teaching assignment process.

A. ACCESS PROCEDURES

The data and manipulative programs associated are held on the IBM 3033 mainframe at the Naval Postgraduate School under user number 1502P. The are contained in an APL workspace called 'OSDATA'. Assuming you have a valid user account on the main NPS computer system, this workspace can be accessed in three steps; linking to user 1502P, accessing his disk, and copying the workspace onto your own disk. Once you are 'logged on' to the system, the command

```
CP link 1502p 195 195 rr
```

will cause a prompt to appear asking for the 'read' password. Enter 'ps3000'. That password allows access to user 1502P's disk. To get the workspace transferred to your own disk, enter

```
acc 195 b/a                          followed by
copy osdata vsaplws b = = a
```

B. RUN PREPARATION INSTRUCTIONS

To effectively run the programs available, data must exist in a six column numerical matrix format coded as shown:

Column Number	Sample Entries	Comments
1	83.1 84.4	This column corresponds to the quarter from which the data is gleaned. '83.1' corresponds to the Fall quarter of the 1982-1983 school year (as in standard fiscal year notation). Similarly, 84.4 corresponds to the summer quarter of the 1983-1984 school year.
2	0 374	This column holds the curriculum code. '0' means that the entries to follow are total number of students predicted and total actually enrolled in the listed course as compared to '374' which means that the figures to follow apply only to curriculum 374.
3	2103 3602	This column holds the OS course number
4	29 0	This column holds the actual number of students attending the course as described by the previous columns. If column two is zero, this number is the total number of students that attended and should correspond to the total listed on the grade record sheet held by the registrar. If this number itself is zero, the course was not taught during the quarter shown in column one.

5	0 15	This column holds the predicted number of students that should attend the course described. If column two is zero, this number is the total number of students that were predicted for the course and should correspond to the totals listed on both the summary and individual course forecasts. If this number itself is zero, either the course was predicted to be empty or the course was not predicted.
6	-1 0 5	This column holds the difference between the actual number of students attending the course and the number predicted (actual minus predicted). Negative numbers imply overestimates, positive numbers imply underestimates.

Prior to any program runs, the goal of your analysis must be defined. The programs are generally divided into two areas, course centered and curricula centered. In the course centered area, the 'COURSEP' program allows a general viewing of the error data for chosen courses. This program displays the data separated into courses ordered by decreasing magnitude of the error term (actual minus predicted; column six) and is used to check quickly the prediction accuracy by course. Then to check for underlying patterns by quarter taught or by individual curricula, the 'COURSE' program can be used. Similarly, in the curricula centered area, the 'CURRICP' program allows a general viewing of the error data separated into individual curricula, and the 'CURRIC' program can be used to further distill the information available.

In summary, two steps must be taken prior to starting program runs:

- Ensure the data exists in the proper format
- Decide what information is needed (define goals)

C. SAMPLE PROGRAM RUNS

To demonstrate effectively the varied uses of the programs available, a hypothetical problem will be created and fully explored. Each program will be discussed in detail as it appears in the problem solution, and both the inputs required and output generated will be shown.

1. Problem Statement

Curriculum 525, Antisubmarine Warfare, and curriculum 374, Air-Ocean Tactical Environmental Support, are going to combine parts of their programs to create a new curriculum: Curriculum 449.5 - Antiwhale Warfare. The condensed course matrix for curriculum 449.5 has OS2103, CS3604, OS3601, and OS3602 in quarters 2, 3, 4 and 5 respectively. As the Operations Analysis Department Chairman, you know that OS2103 and OS3604 can handle the additional students with no problems due to the present size and frequency of those course offerings. Your question, with scheduling coming up for the next fiscal year, is whether the additional 12 students you expect per 'class' will be enough to require the offering of an additional section for OS3601 or OS3602. The figure of twelve students comes from the Air-Ocean curricular office, since curriculum 449.5 will be taught as an adjunct of curriculum 374. These students will be entering in the Fall and spring quarters.

2. Definition of Goals

The first step in the problem solution is to define the information required to answer the question. In this case, we need to know if an additional section is required given an influx of twelve additional students. This problem defines three questions:

- How many students usually take the course in question
- How accurate are those figures
- How good is the estimate of twelve additional students

Since the first items to be considered involve course data, the course related programs will be used. To check later the accuracy of the figure twelve, a curricular comparison can be made with curriculum 374 (since that office is the administrator of the program) using the curricula related programs. Additionally, the data used must be the most recent and complete information available.

3. Course Related Work

The first step we will take is to get a look at the data of interest. The variable 'OS' contains the information we want. The 'COURSEP' program most easily shows this data. Through proper inputs, we can call up the historical performance of OS3601 and OS3602. The 'COURSEP' program provides all the prompts necessary (see figure A.1). The program output is shown in figure A.2 . See figure A.3 for a program listing.

As we can see from figure A.2, the actual number of students that attend each course appears to be somewhere around 15. The predictions are presently perfectly adequate since they show a requirement for only one section in all cases, which holds true. With the addition of twelve students, however, the number of students will be close to 30, necessitating another section. To more adequately show the data, we can look at the 'COURSE' program sorted by curricula. This will break out the summary predictions to better show how well the course is predicted. Figure A.4 shows the inputs, figure A.5 the outputs, and figure A.6 the program itself.

Using only the summary data shown in figure A.5, we can confirm that the actual number of students that attend each course is about 15, and the present one section requirement holds true. We can further observe a possible pattern by quarter (errors appear to increase for the fourth quarter predictions as compared to predictions made for earlier quarters). To more adequately show this pattern in the data, we can look at the 'COURSE' program sorted by quarter. This will break out the predictions to better show quarterly patterns. Figure A.7 shows the inputs, figure A.8 the outputs.

In summary, through use of the course related programs we have found that a course size of approximately fifteen is usual for these courses. Predictions, especially those made for the fourth quarter, tend to overestimate the number of students expected. If the figure of twelve additional students is in fact correct, a new section will probably not be required even though predicted attendance values are in the twenties without the addition of the new curriculum.

```
INPUT        PROMPT

COURSEP OS
                *
   0
                *
3601 3602


  * see COURSEP listing for details
    of the prompts
```

Figure A.1   Input Syntax : COURSEP

```
COURSE : 3601
    ¯83.4      0     3601      14      23     ¯9
    ¯83.4    365     3601       3      11     ¯8
    ¯82.4      0     3601      14      21     ¯7
    ¯82.2      0     3601      14       8      6
    ¯83.4    374     3601       2       5     ¯3
    ¯83.2      0     3601      14      16     ¯2
    ¯83.4    525     3601       9       7      2
    ¯84.2    365     3601       2       0      2
    ¯84.2    525     3601      20      22     ¯2
    ¯83.2    374     3601       4       5     ¯1
    ¯83.2    525     3601      10      11     ¯1
    ¯84.2      0     3601      25      26     ¯1
    ¯84.2    374     3601       3       4     ¯1


COURSE : 3602
    ¯82.1      0     3602       7      11     ¯4
    ¯83.3    374     3602       3       5     ¯2
    ¯84.3    374     3602       2       4     ¯2
    ¯84.3    525     3602      12      10      2
    ¯82.3      0     3602      15      14      1
    ¯83.3      0     3602      15      16     ¯1
    ¯83.3    686     3602       1       0      1
    ¯83.1      0     3602       7       7      0
    ¯83.1    374     3602       1       1      0
    ¯83.1    525     3602       6       6      0
    ¯83.3    525     3602      11      11      0
    ¯84.1      0     3602      12      12      0
    ¯84.1    374     3602       3       3      0
    ¯84.1    525     3602       9       9      0
    ¯84.3      0     3602      14      14      0
```

Figure A.2 Output from Input Sequence shown in Figure A.1 : COURSEP Program.

```
     ∇COURSEP[⎕]∇

    ∇ COURSEP MATRIX;X;I;Y;Z
[1]   ⍝
[2]   ⍝ THIS FUNCTION, GIVEN AN INPUT MATRIX IN THE FOLLOWING FORM :
[3]   ⍝
[4]   ⍝     QTR  CURRIC  CSE  ACTUAL  PREDICTED  (ACTUAL-PREDICTED)
[5]   ⍝
[6]   ⍝ WILL SORT AND DISPLAY THE DATA BY CURRICULUM IN THE
[7]   ⍝ FOLLOWING MANNER :
[8]   ⍝
[9]   ⍝     COURSE
[10]  ⍝
[11]  ⍝          DATA ORDERED BY MAGNITUDE OF ERROR TERM
[12]  ⍝
[13]  ⍝ (ALL QUARTERS WILL BE DISPLAYED AS NEGATIVE NUMBERS. THIS ALLOWS
[14]  ⍝  FOR CORRECT PLACEMENT OF THE OUTPUT COLUMNS WITHOUT RESORTING
[15]  ⍝  TO COMPLEX CHARACTER ARRAY MANIPULATIONS.)
[16]  ⍝
[17]  ⍝ THE PROGRAM GOAL IS TO DISPLAY OBVIOUS PATTERNS THAT MAY THEN BE
[18]  ⍝ LOOKED AT MORE CLOSELY USING THE 'CURRIC' OR 'COURSE' FUNCTIONS.
[19]  ⍝ COURSES MAY BE SELECTED INDIVIDUALLY IF SO DESIRED.
[20]  ⍝
[21]   MATRIX←MATRIX×(⍴MATRIX)⍴ ¯1 1 1 1 1 1
[22]   ⍕'IF YOU WANT ALL COURSES DISPLAYED, ENTER 1'
[23]   ⍕'OTHERWISE, ENTER 0 TO SELECT SOME SUBSET OF THE COURSES'
[24]   →JUMP1×⍳⎕
[25]   ⍕'THE FOLLOWING COURSES ARE AVAILABLE FOR SELECTION :'
[26]   ⍕X←X[⍋X←UNIQUE MATRIX[;3]]
[27]   ⍕'ENTER THE COURSE NUMBERS YOU WANT. ENTER X TO GET THEM ALL.'
[28]   ⍕'(NOTE : YOU MUST CHOOSE MORE THAN ONE COURSE)'
[29]   Y←⍴X←⎕
[30]   →JUMP2
[31]  JUMP1:Y←⍴X←X[⍋X←UNIQUE MATRIX[;3]]
[32]  JUMP2:I←1
[33]  CRSLOOP:⍕' '
[34]   →0×⍳I>Y
[35]   (⍕'COURSE : '),⍕X[I]
[36]   Z←(X[I]=MATRIX[;3])/[1] MATRIX
[37]   ⍕Z ORDER 6
[38]   ⍕' '
[39]   I←I+1
[40]   ⍕' '
[41]   →CRSLOOP
    ∇
    .
```

Figure A.3 'COURSEP' Program

```
INPUT           PROMPT

COURSE OS
                   *
  0
                   *
3601 3602
                   *
  0
                   *
  n


 * see COURSE listing for details
   of the prompts
```

Figure A.4   Input Syntax : COURSE

```
COURSE : 3601
   ¯83.4      0     3601       14        23       ¯9
   ¯82.4      0     3601       14        21       ¯7
   ¯82.2      0     3601       14         8        6
   ¯83.2      0     3601       14        16       ¯2
   ¯84.2      0     3601       25        26       ¯1

   ¯83.4    365     3601        3        11       ¯8
   ¯84.2    365     3601        2         0        2

   ¯83.4    374     3601        2         5       ¯3
   ¯83.2    374     3601        4         5       ¯1
   ¯84.2    374     3601        3         4       ¯1

   ¯83.4    525     3601        9         7        2
   ¯84.2    525     3601       20        22       ¯2
   ¯83.2    525     3601       10        11       ¯1

  GROUP      MEAN    VARIANCE   N   (OF ERROR TERM)
  365.0     ¯3.0     50.0      2.0
  374.0     ¯1.7      1.3      3.0
  525.0      ¯.3      4.3      3.0

OVERALL:     ¯1.5    10.0     8.0


COURSE : 3602
   ¯82.1      0     3602        7        11       ¯4
   ¯82.3      0     3602       15        14        1
   ¯83.3      0     3602       15        16       ¯1
   ¯83.1      0     3602        7         7        0
   ¯84.1      0     3602       12        12        0
   ¯84.3      0     3602       14        14        0

   ¯83.3    374     3602        3         5       ¯2
   ¯84.3    374     3602        2         4       ¯2
   ¯83.1    374     3602        1         1        0
   ¯84.1    374     3602        3         3        0

   ¯84.3    525     3602       12        10        2
   ¯83.1    525     3602        6         6        0
   ¯83.3    525     3602       11        11        0
   ¯84.1    525     3602        9         9        0

   ¯83.3    686     3602        1         0        1

  GROUP      MEAN    VARIANCE   N   (OF ERROR TERM)
  374.0     ¯1.0      1.3      4.0
  525.0       .5      1.0      4.0
  686.0      1.0      1.0      1.0

OVERALL:      ¯.1     1.6     9.0
```

Figure A.5 Output from Input Sequence shown in Figure A.4 : COURSE Program.

```
       ∇COURSE[⎕]∇

      ∇ COURSE MATRIX;X;I;Y;Z;MC;VC;J;X2;Y2;MQ;VQ;W;QTR;NQ;NC;NS;NE;IND;AE;PE
[1]    ⍝
[2]    ⍝ THIS FUNCTION, GIVEN AN INPUT MATRIX IN THE FOLLOWING FORM :
[3]    ⍝
[4]    ⍝     QTR  CURRIC  CSE  ACTUAL  PREDICTED  (ACTUAL-PREDICTED)
[5]    ⍝
[6]    ⍝ WILL SORT AND DISPLAY THE DATA BY CURRICULUM IN THE
[7]    ⍝ FOLLOWING MANNER :
[8]    ⍝
[9]    ⍝     COURSE
[10]   ⍝             QUARTER 1 DATA       CURRICULUM 1 DATA
[11]   ⍝             QUARTER 2 DATA   OR  CURRICULUM 2 DATA
[12]   ⍝             QUARTER 3 DATA                .
[13]   ⍝             QUARTER 4 DATA       CURRICULUM N DATA
[14]   ⍝     MATRIX OF GROUP MEANS, VARIANCES AND N
[15]   ⍝     OVERALL MEAN, VARIANCE AND N
[16]   ⍝
[17]   ⍝ A MATRIX OF ACCURACY, PREDICTIVE AND NO ERROR PERCENTAGES IS
[18]   ⍝ AVAILABLE IF DESIRED.
[19]   ⍝
[20]   ⍝ (ALL QUARTERS WILL BE DISPLAYED AS NEGATIVE NUMBERS. THIS ALLOWS
[21]   ⍝  FOR CORRECT PLACEMENT OF THE OUTPUT COLUMNS WITHOUT RESORTING
[22]   ⍝  TO COMPLEX CHARACTER ARRAY MANIPULATIONS.)
[23]   ⍝
[24]   ⍝ ******** NOTE : THE MEANS AND VARIANCES COMPUTED APPLY   ********
[25]   ⍝ ********        ONLY TO THE INDIVIDUAL CURRICULA !!!!!   ********
[26]   ⍝ ********        ANY SUMMARY DATA (→COLUMN 2 ENTRY = 0)   ********
[27]   ⍝ ********        IS COMPRESSED OUT OF THE COMPUTATION.    ********
[28]   ⍝ ********        THIS CAN BE CHANGED BY REPLACING THE ≠   ********
[29]   ⍝ ********        SIGN IN THE LOOPS (WHERE Z OR W [;2] IS ********
[30]   ⍝ ********        COMPRESSING COLUMN 6 OR WHERE X2 IS      ********
[31]   ⍝ ********        COMPRESSING THE OUTPUT) WITH '=', THIS********
[32]   ⍝ ********        FUNCTION WILL THEN DISPLAY ONLY THE      ********
[33]   ⍝ ********        SUMMARY DATA.                            ********
[34]   ⍝
[35]   ⍝ COURSES MAY BE SELECTED INDIVIDUALLY IF SO DESIRED.
[36]   ⍝
[37]    MATRIX←MATRIX×(⍴MATRIX)⍴ ¯1 1 1 1 1 1
[38]    ⍕'IF YOU WANT ALL COURSES DISPLAYED, ENTER 1'
[39]    ⍕'OTHERWISE, ENTER 0 TO SELECT SOME SUBSET OF THE COURSES'
[40]    →JUMP1×⍳⎕
[41]    ⍕'THE FOLLOWING COURSES ARE AVAILABLE FOR SELECTION :'
[42]    ⍕X←X[⍋X←UNIQUE MATRIX[;3]]
[43]    ⍕'ENTER THE COURSE NUMBERS YOU WANT. ENTER X TO GET THEM ALL.'
[44]    ⍕'(NOTE : YOU MUST CHOOSE MORE THAN ONE COURSE)'
[45]    Y←⍴X←⎕
[46]    →JUMP2
[47]   JUMP1:Y←⍴X←X[⍋X←UNIQUE MATRIX[;3]]
[48]   JUMP2:I←1
[49]    ⍕'THE DATA MAY BE DISPLAYED (TO LOOK AT PATTERNS) IN GROUPINGS BY'
[50]    ⍕'EITHER QUARTER OR CURRIC.  FOR QUARTER GROUPINGS, ENTER 1 ; FOR'
[51]    ⍕'CURRIC GROUPINGS, ENTER 0'
[52]    V←⎕
[53]    ⍕'ACCURACY ERROR, PREDICTIVE ERROR, AND NO ERROR PERCENTAGES ARE'
[54]    ⍕'AVAILABLE. ENTER Y IF DESIRED, N OTHERWISE'
[55]    IND←⎕
[56]   CRSLOOP:⍕' '
[57]    →0×⍳I>Y
[58]    (⍕'COURSE : '),⍕X[I]
[59]    Z←(X[I]=MATRIX[;3])/[1] MATRIX
```

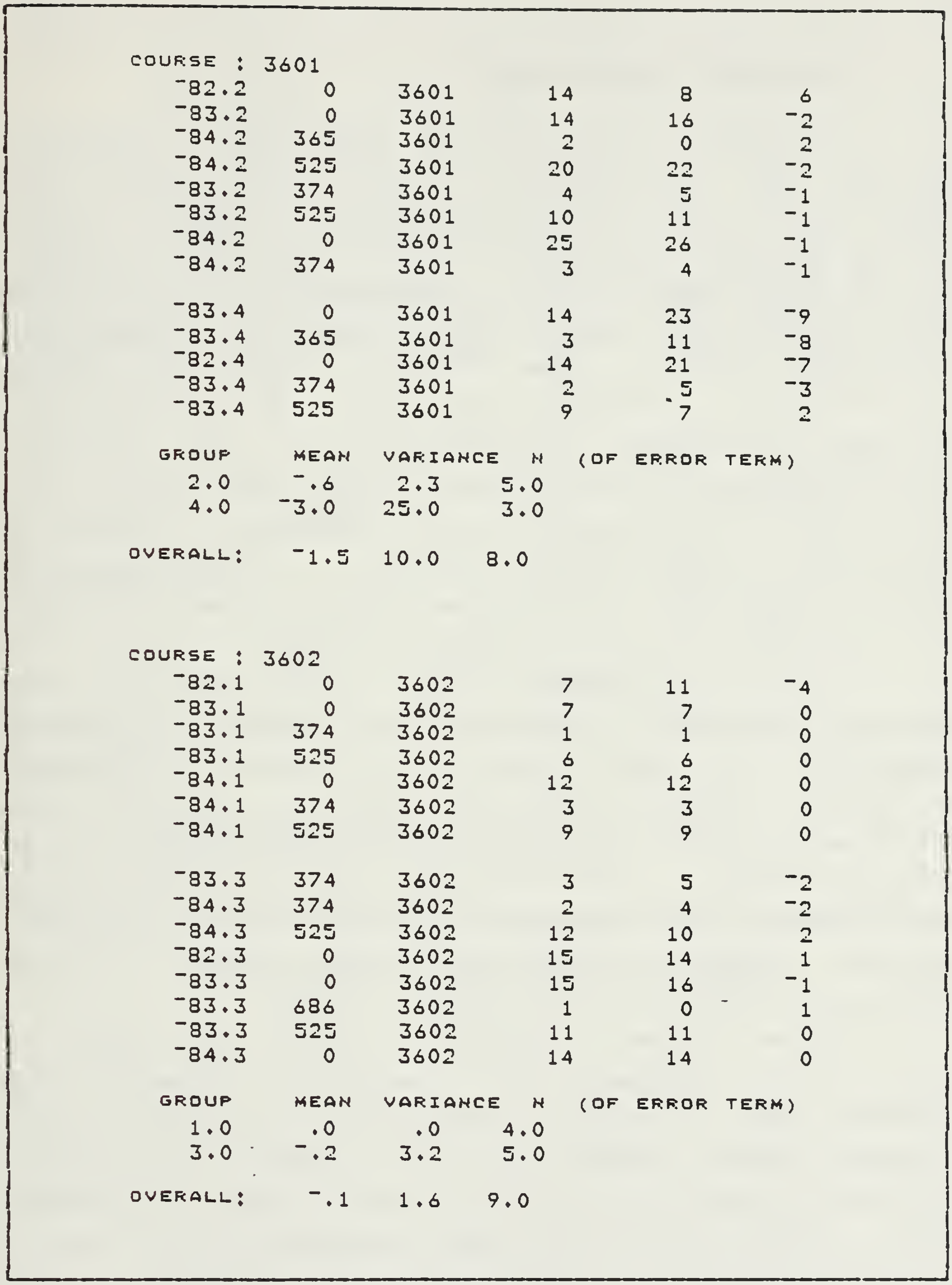
Figure A.6 'COURSE' Program

```
 MC←MEAN(Z[;2]≠0)/Z[;6]
 VC←VARIA(Z[;2]≠0)/Z[;6]
 NC←⍴(Z[;2]≠0)/Z[;6]
 J←1
 →NEXT×⍳V≠1
 QTR←10×(-Z[;1])-⌊-Z[;1]
 →OUT
NEXT:QTR←Z[;2]
OUT:Y2←⍴X2←X2[⍋X2←UNIQUE QTR]
 NS←AE←PE←NQ←MQ←VQ←⍳0
QTRLOOP:→END×⍳J>Y2
 W←((X2[J]=QTR)/[1] Z) ORDER 6
 MQ←MQ,MEAN(W[;2]≠0)/W[;6]
 VQ←VQ,VARIA(W[;2]≠0)/W[;6]
 NQ←NQ,NE←⍴(W[;2]≠0)/W[;6]
 →JUMP3×⍳IND='N'
 PE←PE,(+/(W[;2]≠0)/(W[;4]=0)∨W[;5]=0)÷NE
 AE←AE,(+/(W[;2]≠0)/(W[;4]≠0)∧(W[;5]≠0)∧W[;6]≠0)÷NE
 NS←NS,(+/(W[;2]≠0)/(W[;4]≠0)∧(W[;5]≠0)∧W[;6]=0)÷NE
JUMP3:⍕W
 ⍕' '
 J←J+1
 →QTRLOOP
END:I←I+1
 ⍕'  GROUP    MEAN  VARIANCE  N  (OF ERROR TERM)'
 7 1 ⍕(X2≠0)/[1]((⍉X2,[0.5] MQ),[2] VQ),[2] NQ
 ⍕' '
 (⍕'OVERALL: '), 6 1 ⍕MC,VC,NC
 ⍕' '
 →JUMP4×⍳IND='N'
 ⍕'  ERROR PERCENTAGES'
 ⍕'  G̲R̲O̲U̲P̲    P̲R̲E̲    A̲C̲C̲   N̲O̲N̲E̲    N̲'
 7 2 ⍕(X2≠0)/[1](((⍉X2,[0.5] PE),[2] AE),[2] NS),[2] NQ
 ⍕' '
JUMP4:⍕' '
 →CRSLOOP
 ∇
```

Figure A.6 (continued)

```
INPUT         PROMPT

COURSE OS
                 *
   0
                 *
3601 3602
                 *
   1
                 *
   n


  * see COURSE listing for details
    of the prompts
```

Figure A.7 Input Syntax : COURSE



```
COURSE : 3601
   ¯82.2     0    3601      14       8       6
   ¯83.2     0    3601      14      16      ¯2
   ¯84.2   365    3601       2       0       2
   ¯84.2   525    3601      20      22      ¯2
   ¯83.2   374    3601       4       5      ¯1
   ¯83.2   525    3601      10      11      ¯1
   ¯84.2     0    3601      25      26      ¯1
   ¯84.2   374    3601       3       4      ¯1

   ¯83.4     0    3601      14      23      ¯9
   ¯83.4   365    3601       3      11      ¯8
   ¯82.4     0    3601      14      21      ¯7
   ¯83.4   374    3601       2       5      ¯3
   ¯83.4   525    3601       9       7       2

  GROUP      MEAN   VARIANCE  N  (OF ERROR TERM)
   2.0       ¯.6     2.3     5.0
   4.0      ¯3.0    25.0     3.0

OVERALL:    ¯1.5    10.0    8.0


COURSE : 3602
   ¯82.1     0    3602       7      11      ¯4
   ¯83.1     0    3602       7       7       0
   ¯83.1   374    3602       1       1       0
   ¯83.1   525    3602       6       6       0
   ¯84.1     0    3602      12      12       0
   ¯84.1   374    3602       3       3       0
   ¯84.1   525    3602       9       9       0

   ¯83.3   374    3602       3       5      ¯2
   ¯84.3   374    3602       2       4      ¯2
   ¯84.3   525    3602      12      10       2
   ¯82.3     0    3602      15      14       1
   ¯83.3     0    3602      15      16      ¯1
   ¯83.3   686    3602       1       0       1
   ¯83.3   525    3602      11      11       0
   ¯84.3     0    3602      14      14       0

  GROUP      MEAN   VARIANCE  N  (OF ERROR TERM)
   1.0        .0       .0    4.0
   3.0       ¯.2      3.2    5.0

OVERALL:     ¯.1     1.6    9.0
```

Figure A.8 Output from Input Sequence shown in Figure A.7 : COURSE Program.

4. Curricula Related Work

In looking at the curricula information, we want to find out just how good the figure 'twelve' is for the expected input from the Antiwhale Warfare curriculum. Since curriculum 374 will be administering the program, we will use the data from 374 as an approximation of curriculum 449.5's performance. Again, the variable 'OS' contains the information we want. The 'CURRICP' program most easily shows this data. Through proper input, we can call up the historical performance of curriculum 374. The 'CURRICP' program provides all the prompts necessary (see figure A.9). Since we are only interested in one curriculum but two are required for proper program execution, a dummy curriculum (333) is requested. The program output is shown in figure A.10 See figure A.11 for a program listing.

As we can see from figure A.10, curriculum 374 routinely appears to overestimate the number of students expected for any given course. To show more adequately patterns of concern in the data (such as performance for OS3601 and OS3602), we can look at the 'CURRIC' program sorted by course. This will break out the predictions to better show how well curriculum 374 performs for individual courses. Figure A.12 shows the inputs, figure A.13 the outputs, and figure A.14 the program itself.

As we can now see from the course data shown in figure A.13, curriculum 374 generally overstates the expected number of students attending courses OS3601 and, to a lesser extent, OS3602. We can't observe any possible patterns by quarter due to the limited number of observations. A brief look at the error percentage table shows no predictive errors for either course (they're taught or not taught when they say so - no 'zeros' in the actual or predicted columns), and an improvement in accuracy from 3601 to 3602 (0% error free to 50% error free respectively).

In summary, through use of the curricula related programs we have found that a slight overestimation of students expected is usual for these courses. This trend may or may not hold for increased class sizes (of, say, twelve students vice three or four). Therefor, combined with the information we previously obtained from the course related programs, we conclude that new sections will probably not be required.

```
INPUT          PROMPT

CURRICP OS

                  *

  0

                  *

 374 333


   * see COURSEP listing for details
     of the prompts
```

Figure A.9 Input Syntax : CURRICP

```
CURRICULUM : 374
    ¯83.4    374     2101          0          6         ¯6
    ¯83.4    374     2103          2          6         ¯4
    ¯83.2    374     2103          1          4         ¯3
    ¯83.3    374     3604          1          4         ¯3
    ¯83.4    374     3601          2          5         ¯3
    ¯83.3    374     3602          3          5         ¯2
    ¯83.4    374     3603          4          6         ¯2
    ¯84.1    374     3604          2          4         ¯2
    ¯84.3    374     3602          2          4         ¯2
    ¯83.2    374     3601          4          5         ¯1
    ¯83.2    374     3604          0          1         ¯1
    ¯84.1    374     2103          0          1         ¯1
    ¯84.1    374     3401          0          1         ¯1
    ¯84.2    374     2103          2          3         ¯1
    ¯84.2    374     3601          3          4         ¯1
    ¯84.3    374     3604          2          3         ¯1
    ¯83.1    374     3602          1          1          0
    ¯83.1    374     3604          4          4          0
    ¯84.1    374     3602          3          3          0
    ¯84.1    374     3603          1          1          0


CURRICULUM : 333
```

Figure A.10   Output from Input Sequence shown in Figure A.9 : CURRICP Program.

```
      ∇CURRICP[⎕]∇

     ∇ CURRICP MATRIX;X;I;Y;Z
[1]   ⍝
[2]   ⍝ THIS FUNCTION, GIVEN AN INPUT MATRIX IN THE FOLLOWING FORM :
[3]   ⍝
[4]   ⍝     QTR   CURRIC   CSE   ACTUAL   PREDICTED  (ACTUAL-PREDICTED)
[5]   ⍝
[6]   ⍝ WILL SORT AND DISPLAY THE DATA BY CURRICULUM IN THE
[7]   ⍝ FOLLOWING MANNER :
[8]   ⍝
[9]   ⍝     CURRICULUM
[10]  ⍝
[11]  ⍝            DATA ORDERED BY MAGNITUDE OF ERROR TERM
[12]  ⍝
[13]  ⍝ (ALL QUARTERS WILL BE DISPLAYED AS NEGATIVE NUMBERS. THIS ALLOWS
[14]  ⍝  FOR CORRECT PLACEMENT OF THE OUTPUT COLUMNS WITHOUT RESORTING
[15]  ⍝  TO COMPLEX CHARACTER ARRAY MANIPULATIONS.)
[16]  ⍝
[17]  ⍝ THE GOAL OF THIS PROGRAM IS TO DISPLAY OBVIOUS PATTERNS THAT
[18]  ⍝ MAY THEN BE LOOKED AT MORE CLOSELY USING THE 'COURSE' OR 'CURRIC'
[19]  ⍝ FUNCTIONS. CURRICULA MAY BE SELECTED INDIVIDUALLY IF SO DESIRED.
[20]  ⍝
[21]   MATRIX←MATRIX×(⍴MATRIX)⍴ ¯1 1 1 1 1 1
[22]   ⍕'IF YOU WANT ALL CURRICULA DISPLAYED, ENTER 1'
[23]   ⍕'OTHERWISE, ENTER 0 TO SELECT SOME SUBSET OF THE CURRICULA'
[24]   →JUMP1×⍳⎕
[25]   ⍕'THE FOLLOWING CURRICULA ARE AVAILABLE FOR SELECTION :'
[26]   ⍕X←X[⍋X←UNIQUE MATRIX[;2]]
[27]   ⍕'ENTER THE CURRIC NUMBERS YOU WANT. ENTER X TO GET THEM ALL.'
[28]   ⍕'(NOTE : YOU MUST ENTER MORE THAN ONE CURRICULUM)'
[29]   Y←⍴X←⎕
[30]   →JUMP2
[31]  JUMP1:Y←⍴X←X[⍋X←UNIQUE MATRIX[;2]]
[32]  JUMP2:I←1
[33]  CRCLOOP:⍕' '
[34]   →0×⍳I>Y
[35]   (⍕'CURRICULUM : '),⍕X[I]
[36]   Z←(X[I]=MATRIX[;2])/[1] MATRIX
[37]   ⍕Z ORDER 6
[38]   I←I+1
[39]   ⍕' '
[40]   ⍕' '
[41]   →CRCLOOP
      ∇
```

Figure A.11 'CURRICP' Program

```
INPUT          PROMPT

CURRIC OS
                  *

  0
                  *

 374 333
                  *

  3
                  *

  y


  * see COURSE listing for details
    of the prompts
```

Figure A.12   Input Syntax : CURRIC

```
CURRICULUM : 374
   ¯83.4    374    2101        0        6       ¯6

   ¯83.4    374    2103        2        6       ¯4
   ¯83.2    374    2103        1        4       ¯3
   ¯84.1    374    2103        0        1       ¯1
   ¯84.2    374    2103        2        3       ¯1

   ¯84.1    374    3401        0        1       ¯1

   ¯83.4    374    3601        2        5       ¯3
   ¯83.2    374    3601        4        5       ¯1
   ¯84.2    374    3601        3        4       ¯1

   ¯83.3    374    3602        3        5       ¯2
   ¯84.3    374    3602        2        4       ¯2
   ¯83.1    374    3602        1        1        0
   ¯84.1    374    3602        3        3        0

   ¯83.4    374    3603        4        6       ¯2
   ¯84.1    374    3603        1        1        0

   ¯83.3    374    3604        1        4       ¯3
   ¯84.1    374    3604        2        4       ¯2
   ¯83.2    374    3604        0        1       ¯1
   ¯84.3    374    3604        2        3       ¯1
   ¯83.1    374    3604        4        4        0
  GROUP      MEAN   VARIANCE   N   (OF ERROR TERM)
 2101.0     ¯6.0     1.0     1.0
 2103.0     ¯2.3     2.3     4.0
 3401.0     ¯1.0     1.0     1.0
 3601.0     ¯1.7     1.3     3.0
 3602.0     ¯1.0     1.3     4.0
 3603.0     ¯1.0     2.0     2.0
 3604.0     ¯1.4     1.3     5.0

OVERALL :     ¯1.7    2.3   20.0

  ERROR PERCENTAGES
  GROUP     PRE      ACC      NONE     N
2101.00    1.00      .00      .00    1.00
2103.00     .25      .75      .00    4.00
3401.00    1.00      .00      .00    1.00
3601.00     .00     1.00      .00    3.00
3602.00     .00      .50      .50    4.00
3603.00     .00      .50      .50    2.00
3604.00     .20      .60      .20    5.00


CURRICULUM : 333
  GROUP      MEAN   VARIANCE   N   (OF ERROR TERM)


OVERALL :     1.0     .0      .0

  ERROR PERCENTAGES
  GROUP     PRE      ACC      NONE     N
```

Figure A.13 Output from Input Sequence shown in Figure A.12 : CURRIC Program.

```
      ∇CURRIC[⎕]∇

    ∇ CURRIC MATRIX;X;I;Y;Z;MC;VC;J;X2;Y2;MQ;VQ;W;A;NS;NE;IND;AE;PE;V;NC;NQ
[1]   ⍝
[2]   ⍝ THIS FUNCTION, GIVEN AN INPUT MATRIX IN THE FOLLOWING FORM :
[3]   ⍝
[4]   ⍝     QTR  CURRIC  CSE  ACTUAL  PREDICTED  (ACTUAL-PREDICTED)
[5]   ⍝
[6]   ⍝ WILL SORT AND DISPLAY THE DATA BY CURRICULUM IN THE
[7]   ⍝ FOLLOWING MANNER :
[8]   ⍝
[9]   ⍝     CURRICULUM
[10]  ⍝                         QUARTER 1 DATA       COURSE 1 DATA
[11]  ⍝                         QUARTER 2 DATA OR   COURSE 2 DATA
[12]  ⍝                         QUARTER 3 DATA               .
[13]  ⍝                         QUARTER 4 DATA       COURSE N DATA
[14]  ⍝            MATRIX OF GROUP MEANS, VARIANCES AND N
[15]  ⍝            OVERALL MEAN, VARIANCE AND N
[16]  ⍝
[17]  ⍝ A MATRIX OF ACCURACY, PREDICTIVE AND NO ERROR PERCENTAGES IS
[18]  ⍝ AVAILABLE IF DESIRED.
[19]  ⍝
[20]  ⍝ (ALL QUARTERS WILL BE DISPLAYED AS NEGATIVE NUMBERS. THIS ALLOWS
[21]  ⍝  FOR CORRECT PLACEMENT OF THE OUTPUT COLUMNS WITHOUT RESORTING
[22]  ⍝  TO COMPLEX CHARACTER ARRAY MANIPULATIONS.)
[23]  ⍝
[24]  ⍝ CURRICULA MAY BE SELECTED INDIVIDUALLY IF SO DESIRED.
[25]  ⍝
[26]   MATRIX←MATRIX×(⍴MATRIX)⍴ ¯1 1 1 1 1 1
[27]   ⍕'IF YOU WANT ALL CURRICULA DISPLAYED, ENTER 1'
[28]   ⍕'OTHERWISE, ENTER 0 TO SELECT SOME SUBSET OF THE CURRICULA'
[29]   →JUMP1×⍳⎕
[30]   ⍕'THE FOLLOWING CURRICULA ARE AVAILABLE FOR SELECTION :'
[31]   ⍕X←X[⍋X←UNIQUE MATRIX[;2]]
[32]   ⍕'ENTER THE CURRIC NUMBERS YOU WANT. ENTER X TO GET THEM ALL.'
[33]   ⍕'(NOTE : YOU MUST ENTER MORE THAN ONE CURRICULUM)'
[34]   Y←⍴X←⎕
[35]   →JUMP2
[36]  JUMP1:Y←⍴X←X[⍋X←UNIQUE MATRIX[;2]]
[37]  JUMP2:I←1
[38]   ⍕'THE DATA MAY BE DISPLAYED (TO LOOK AT PATTERNS) IN GROUPINGS BY'
[39]   ⍕'EITHER QUARTER OR COURSE.  FOR QUARTER GROUPINGS, ENTER 1 ; FOR'
[40]   ⍕'COURSE GROUPINGS, ENTER 3.'
[41]   V←⎕
[42]   ⍕'ACCURACY ERROR, PREDICTIVE ERROR AND NO ERROR PERCENTAGE MATRIX?'
[43]   ⍕'ENTER Y OR N'
[44]   IND←⍞
[45]  CPCLOOP:⍕' '
[46]   →0×⍳I>Y
[47]   (⍕'CURRICULUM : '),⍕X[I]
[48]   Z←(X[I]=MATRIX[;2])/[1] MATRIX
[49]   MC←MEAN Z[;6]
[50]   VC←VARIA Z[;6]
[51]   NC←⍴Z[;6]
[52]   J←1
[53]   →NEXT×⍳V=3
[54]   A←10×(-Z[;V])-⌊-Z[;V]
[55]   →OUT
[56]  NEXT:A←Z[;V]
[57]  OUT:Y2←⍴X2←X2[⍋X2←UNIQUE A]
[58]   NS←AE←PE←NQ←MQ←VQ←⍳0
[59]  QTRLOOP:→END×⍳J>Y2
```

Figure A.14 'CURRIC' Program

```
[60]  W←((X2[J]=A)/[1] Z) ORDER 6
[61]  MQ←MQ,MEAN W[;6]
[62]  VQ←VQ,VARIA W[;6]
[63]  NQ←NQ,NE←⍴W[;6]
[64]  →JUMP3×⍳IND='N'
[65]  PE←FE,(+/(W[;4]=0)∨W[;5]=0)÷NE
[66]  AE←AE,(+/(W[;4]≠0)∧(W[;5]≠0)∧W[;6]≠0)÷NE
[67]  NS←NS,(+/(W[;4]≠0)∧(W[;5]≠0)∧W[;6]=0)÷NE
[68] JUMP3:⍕W
[69]  ⍕' '
[70]  J←J+1
[71]  →QTRLOOP
[72] END:I←I+1
[73]  ⍕'  GROUP    MEAN  VARIANCE  N  (OF ERROR TERM)'
[74]  7 1 ⍕((⍉X2,[0.5] MQ),[2] VQ),[2] NQ
[75]  ⍕' '
[76]  (⍕'OVERALL : '), 6 1 ⍕MC,VC,NC
[77]  ⍕' '
[78]  →JUMP4×⍳IND='N'
[79]  ⍕'  ERROR PERCENTAGES'
[80]  ⍕'  GROUP    PRE    ACC   NONE    N'
[81]  7 2 ⍕(((⍉X2,[0.5] PE),[2] AE),[2] NS),[2] NQ
[82]  ⍕' '
[83] JUMP4:⍕' '
[84]  →CRCLOOP
      ∇
      ◆
```

Figure A.14 (continued)

D. DATA COLLECTION AND MAINTENANCE

To provide the matrix of historical data to be displayed and analyzed, two sources are used. The prediction data comes from the course loading forecast delivered to the department chairmen (see figure A.15). The easiest way to compile that data is to obtain a copy of the individual course predictions and, on the forecast itself, change the letter codes found in the "S/D/T" column into curricula numbers (using the coding list contained in appendix B) with associated totals. For example, the totals you would write in from figure A.15 for OS3602 are 525: 9, 374: 3, and 0 (total): 12. The other document required is the course grade sheet held by the registrar (see figure A.16). This information is subject to the privacy act and requires a written request for permission from the department chairman involved to the office of the registrar for access to the files. The records themselves are kept according to academic year. This notation is confusing; to prevent problems, ask for something like the '83-84 school year' data. These records are not allowed to leave the office, and at last check were not available on the computer. The data must be recorded by hand. A convenient format is:

Quarter Course Section Total# #Attending/Curricula

Once the above listed data have been collected and are in front of you and a computer terminal, a CMS file as shown in figure A.17 can be created by combining the information. The columns correspond to the quarter, curriculum, course, actual and predicted number of students. The CMS file thus created can be transferred to APL and named, for instance, CSNEXT through the following command sequence:

```
(get into APL)

)LOAD 990 CMSIO
```

```
)WSID HOLD
OSNEXT <= CMSREAD
```

(follow prompts to read your CMS file)

```
)ERASE CMSIO
)SAVE
```

NOTES : 1) The APL variable 'OSNEXT' now exists on your disk in the APL workspace 'HOLD'

2) The '<=' character shown corresponds to the left-pointing assignment arrow in APL

To add the (Actual minus Predicted) term to the matrix, enter:

```
CSNEXT <= OSNEXT,[2] OSNEXT[;4] - OSNEXT[;5]
```

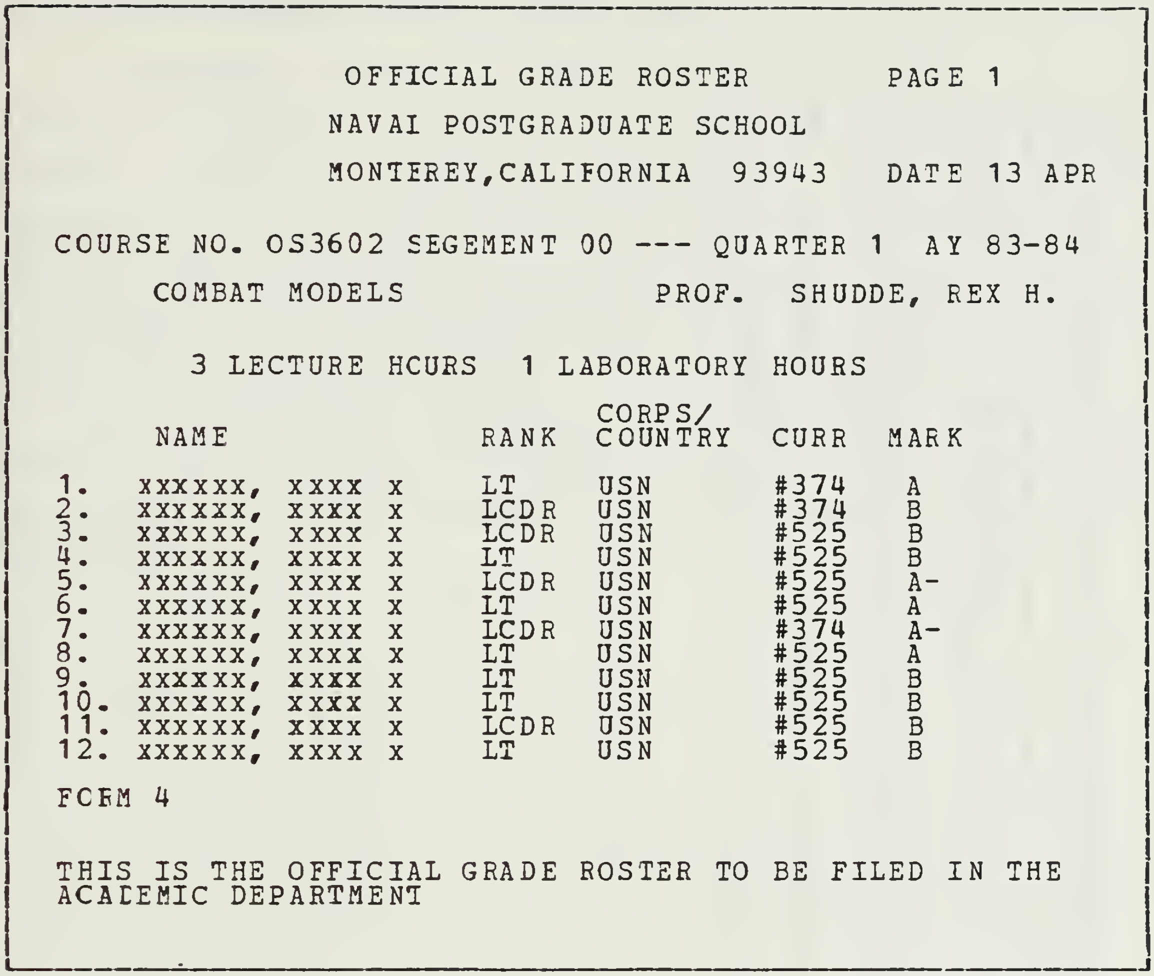
To add this matrix to the variable OS (assuming that OS is in the workspace), enter:

```
CS <= OS,[1] OSNEXT
```

The entire updating procedure should take about one full workday for each year of data.

```
ACADEMIC DEPARTMENT CHAIRMAN REPORT FOR OPERATIONS RESEARCH DEPT  1 AUGUST 1983
ACADEMIC YEAR 84     ACADEMIC QUARTER 1                         FAL83 TRU SUMR84

COURSE LECT/LAB SECTION STUDENTS AV  S/D/T  SEGEMENT INSTRUCTOR  ROOM  TEXTS
*********************************************************************************
OA 2600  4    0   RC3301     5    0  RCMS
                  RG3301     4    0  RGMS
                  RG4101     3    0  RGMS
                C RG4101     5    1  RGMS01
                  RN3301     2    0  RNMS
                C RN4101     5    1  RNMS01
                  RN4102     7    0  RNMS
                C RS4101     2    1  RSMS01

  TOTAL NO. STUDENTS ENROLLED IS 33

OA 2651  4    0 R RA4101     5    1  RAMS01
                  RA4101     6    0  RAMS
                  RC3302     1    0  RCMS

  TOTAL NO. STUDENTS ENROLLED IS 12


    .
    .
    .


OS 3602  4    1   IX2102     2    0  IXMS
                  IX2103     2    0  IXMS
                  IX2104     2    0  IXMS
                  IX2301     1    0  IXMS
                  IX2302     1    0  IXMS
                  IX2303     1    0  IXMS
                  XT2201     1    0  XTMS
                  XT2303     1    0  XTMS
                  XT3101     1    0  XTMS

  TOTAL NO. STUDENTS ENROLLED IS 12
```

Figure A.15 Individual Course Prediction Format



```
                   OFFICIAL GRADE ROSTER        PAGE 1
                  NAVAL POSTGRADUATE SCHOOL
                  MONTEREY,CALIFORNIA  93943    DATE 13 APR

COURSE NO. OS3602 SEGEMENT 00 --- QUARTER 1  AY 83-84
     COMBAT MODELS                PROF.  SHUDDE, REX H.

       3 LECTURE HOURS  1 LABORATORY HOURS

                                    CORPS/
     NAME                    RANK   COUNTRY   CURR   MARK

1.   xxxxxx, xxxx x          LT     USN       #374    A
2.   xxxxxx, xxxx x          LCDR   USN       #374    B
3.   xxxxxx, xxxx x          LCDR   USN       #525    B
4.   xxxxxx, xxxx x          LT     USN       #525    B
5.   xxxxxx, xxxx x          LCDR   USN       #525    A-
6.   xxxxxx, xxxx x          LT     USN       #525    A
7.   xxxxxx, xxxx x          LCDR   USN       #374    A-
8.   xxxxxx, xxxx x          LT     USN       #525    A
9.   xxxxxx, xxxx x          LT     USN       #525    B
10.  xxxxxx, xxxx x          LT     USN       #525    B
11.  xxxxxx, xxxx x          LCDR   USN       #525    B
12.  xxxxxx, xxxx x          LT     USN       #525    B

FORM 4

THIS IS THE OFFICIAL GRADE ROSTER TO BE FILED IN THE
ACADEMIC DEPARTMENT
```

Figure A.16   Sample Grade Sheet (names removed)

```
84.1   0 2103 49 44
84.1 365 2103 29 28
84.1 366 2103 11 15
84.1 374 2103  0  1
84.1 590 2103  4  0
84.1 591 2103  2  0
84.1 535 2103  1  0
84.1 600 2103  2  0
84.1   0 3001 25 22
84.1 368 3001 25 22
84.1   0 3004 28 28
84.1 367 3004 28 28
84.1   0 3006 63 59
84.1 837 3006 21 22
84.1 857 3006 11 10
84.1 819 3006  2  2
84.1 827 3006  2  2
84.1 847 3006 13 13
84.1 815 3006  5  4
84.1 813 3006  1  1
84.1 817 3006  5  1
84.1 360 3006  2  0
84.1 620 3006  1  0
84.1   0 3102 56 68
84.1 620 3102  2  2
84.1 837 3102 22 23
84.1 857 3102  8 10
84.1 819 3102  0  4
84.1 827 3102  2  3
84.1 815 3102  6  7
84.1 817 3102  8  2
84.1 814 3102  6  6
84.1 813 3102  1  3
84.1 847 3102  1  0
84.1   0 3104 29 10
84.1 817 3104  4  1
84.1 570 3104 10  7
84.1 360 3104  1  2
84.1 813 3104  2  0
84.1 819 3104  4  0
```

Figure A.17   CMS File Format

E. OTHER AVAILABLE PROGRAMS AND DATA

A complete listing of the functions available in the 'OSDATA' workspace is shown in table VII . The COURSEP, COURSE, CURRICP and CURRIC programs have been previously demonstrated. The only non-trivial function remaining to be shown is the 'WILCOXSR' program (see figure A.18). As shown in the figure, this program is based on the Wilcoxian Signed Rank Test[11] detailed in [Ref. 5 pp 280-288].

A self-descriptive listing of the variables contained in the workspace is shown in table VIII . The latest data presently available is the spring of 1984 (academic year 83-84) figures.

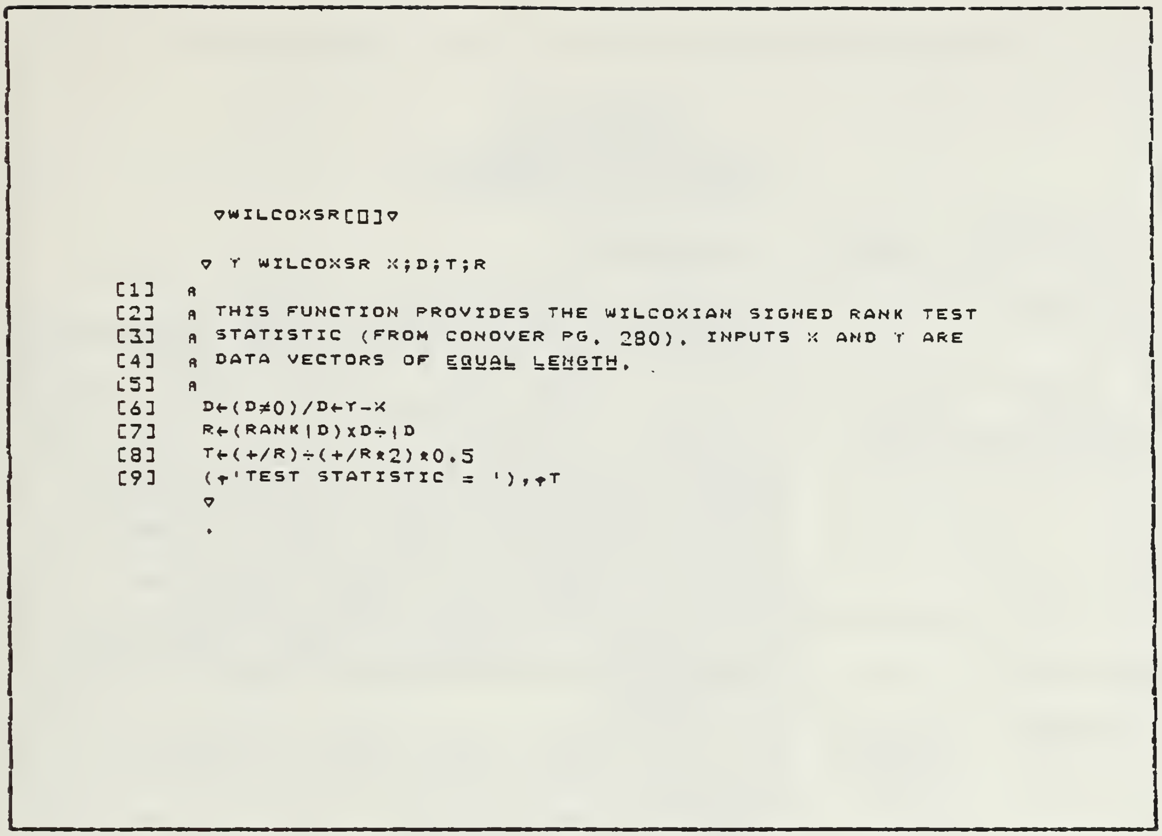
---

[11] For a complete description of the various hypotheses that may be explored, see [Ref. 5 pp 280-288]. In the body of the thesis, the magnitudes of the forecasting errors (y) were tested against the magnitudes of prediction errors derived from course matrices (x). The hypothesis tested (with rejection the desired result) was that the 'x' values tend to be equal to or larger than the 'y' values. This was one of six possible interpretations of the null hypothesis.

TABLE VII

Functions Available

Function	Syntax	Description
********	********	**********************************
COURSEP	COURSEP mtx	(previously shown)
COURSE	COURSE mtx	(previously shown)
CURRICP	CURRICP mtx	(previously shown)
CURRIC	CURRIC mtx	(previously shown)
WILCOXSR	y WILCOXSR x	Shows the Wilcoxian Signed Rank Test results (see description in text)
ORDER	A ORDER B	Orders rows of matrix A according to decreasing magnitude of column B, returning the ordered matrix
UNIQUE	UNIQUE A	Returns the unique elements of the vector or matrix A in ascending order
MEAN	MEAN X	Returns the arithmetic mean of vector X
VARIA	VARIA X	Returns the unbiased estimate of the variance of vector X



```
      ∇WILCOXSR[⎕]∇

     ∇ Y WILCOXSR X;D;T;R
[1]   ⍝
[2]   ⍝ THIS FUNCTION PROVIDES THE WILCOXIAN SIGNED RANK TEST
[3]   ⍝ STATISTIC (FROM CONOVER PG. 280). INPUTS X AND Y ARE
[4]   ⍝ DATA VECTORS OF EQUAL LENGTH.
[5]   ⍝
[6]    D←(D≠0)/D←Y-X
[7]    R←(RANK|D)×D÷|D
[8]    T←(+/R)÷(+/R*2)*0.5
[9]    (⍕'TEST STATISTIC = '),⍕T
     ∇
```

Figure A.18   Wilcoxian Signed Rank Test Program

TABLE VIII

Variables Available

VARIABLE	SHAPE	DESCRIPTION
********	*******	************
OS82	69x6	AY (Academic year) 81-82 data; includes summary information only (no data for individual curricula)
OS83	255x6	AY 82-83 data; complete listing
OSTRK83	140x8	AY 82-83 non-entry course data. The additional columns contained in this matrix (columns 6 and 8) correspond to an 'adjusted prediction' and an 'adjusted error' term, derived from the course matrices. This array was used to check the feasability of making adjustments based solely on the 'track' a student takes using the Wilcoxian Signed Rank Test previously mentioned.
OS84	228x6	AY 83-84 data; only qtrs 1,2 and 3
CS	552x6	Combination of all of above (without repetition)
OA83	113x9	This matrix is set up in a different format from those previously discussed. The nine columns represent Quarter, OA course number, Total attending, USN attending, USA attending, USMC attending, Others attending, Total predicted, and Total attending minus total predicted, respectively. The original intent, dropped due to time constraints, was to use these figures to aid in OA elective scheduling by looking at expected number of upcoming free electives by service through use of the OA course matrix.

APPENDIX B

SCHEDULING CODES

This appendix contains an alphabetical listing of the specialty codes used in the scheduling process as of August 1984. This listing only uses the first two letters of any individual coding because the remaining letters specify either the degree awarded or the degree and month of input, neither of which change the curriculum associated. Using these codes, numbers of students predicted by any given curriculum for any given course can be taken from the individual course forecast promulgated by the programming office. Code 0301 at extension 2773 at the Naval Postgraduate School can be contacted for updated information.

Curricula Numbers	Specialty Degree Codes	Program
610	AC-- AT--	Aero Engr, incl Engr Sci
611	AX--	Aeroelectronics
368	CS--	Computor Science
600	DC--	Electrical Engineering (Comms) Note: DCTR is curriculum 460
590	EC-- ED-- EG-- ER--	Electrical Engineering ; various sub-specialties. Note: EETR is curriculum 460
380	GM--	Math option of advanced science

Curricula Numbers	Specialty Degree Codes	Program
381	GP--	Physics option of adv. science
620	HC-- HM--	Communications Management
525	IX--	Antisubmarine Warfare Note: IETR is curriculum 460
365	JM-- JN-- JO-- JP-- JS--	Command, Control and Communications ; various sub-specialties Note: JCTR and JXTR are from curricula 460
595	KE--	Electronic Warfare Systems Technology Note: KETR is curric 460
817	ME--	Admin Sci - Econ, USMC
837	MF--	Admin Sci - Financial mgt
857	MH--	Admin Sci - Human Resource mgt
819	MI--	Admin Sci - Inventory
817	MK--	Admin Sci - International stu.
827	MM--	Admin Sci - Material
817	MO--	Admin Sci - Ops Research, Army
847	MP--	Admin Sci - Personnel
815	MR--	Admin Sci - Procurement
817	MS--	Admin Sci - MGMT Sci, Coast Guard
814	MT--	Admin Sci - Transportation

Curricula Numbers	Specialty Degree Codes	Program
813	MV--	Admin Sci - Material Movement
817	MX--	Admin Sci - Civ. Federal Program
570	NE--	Mechanical Engineering
	NF--	ME via 460
	NI--	ME via 460
	NR--	ME via 460
	NT--	ME via 460
441	OH--	Hydrographic Science
440	OS--	Oceanography Note: OCTR is 460
367	PL--	Information Systems (from CS)
360	RA--	Ops Analysis, Army
460	RC--	Engr Sci for Ops Analysis
360	RG--	Ops Analysis, General
	RN--	Navy
	RS--	Supply
	RP--	PHD program
591	SE--	Space Systems Engineering
366	SO--	Space Systems Operations Note: SXTR and SOTR are curriculum 460
535	UX--	Underwater Acoustics Note: UETR is curriculum 460

Curricula Numbers	Specialty Degree Codes	Program
530	WA--	Weapons Systems Technology
	WC--	
	WE--	
	WM--	
532	WN--	Weapons Systems Science
530	WP--	Weapons Systems Technology
531	WS--	Weapons Systems Science
372	XA--	Meteorology
373	XS--	Met. and Oceanography
374	XT--	Air-Ocean Tactical Env. Support
		Note: XOTR and XXTR are 460
683	YA--	NSA - USSR, Europe; USN
682	YB--	NSA - Asia/Pacific; USN
681	YC--	NSA - Middle East, Africa; USN
	YD--	- S. Asia; USAF, USA
682	YE--	NSA - Asia/Pacific; USAF, USA
681	YF--	NSA - S. Africa; USAF, USA
	YG--	- N. Africa; USAF, USA
	YM--	- Middle East; USAF, USA
687	YN--	NSA - Nuclr/Plng; USN
684	YO--	NSA - Intl/Org/Neg; USN

Curricula Numbers	Specialty Degree Codes	Program
686	YP--	NSA - Strat/Plng; USN
683	YR--	NSA - USSR/E. Europe
825	YS--	NSA - Navy Intell; USN
683	YW--	NSA - USSR/W. Europe

APPENDIX C
DETAILED SUMMARY OF RESULTS

This appendix contains a brief description of the results found for each course and curriculum. These results are listed in ascending numerical order first for the courses then the curricula. Courses or curricula that cause problems are flagged with a 'P' in front of the identifying number.

Course Number	Comments
******	***
P 2101	Not taught, only predicted.
P 2102	Large magnitude errors, generally underestimates. Curriculum 590 causes most of the problems, curriculum 600 to a lesser extent. No quarterly pattern.
P 2103	Generally good except for curriculum 531 which never sends people but has predicted 40, 27 and 7. No significant quarterly pattern.
P 3001	Prediction errors in third and fourth quarters - not predicted but taught. Only curriculum 368.
3002	Generally good. Only taken by curriculum 825 though others are predicted. Need more data.
3003	Generally good. Need more data.
3004	Generally good.
3005	Generally good. Need more data.

Course Number	Comments
3006	Generally good except for curriculum 817. Only taught first and third quarters. Overall class sizes fairly stable at 63 and 87 respectively (three sections required).
P 3008	Definite prediction problem that appears to be improving. Curriculum 365 provides most of the input.
P 3101	Curriculum 460 not predicted generally causing a significant underestimation (usually results in adding an additional section). More data needed.
P 3102	Overall significant overestimation from cumulative effect of 8XX curricula. Worst in third quarter.
P 3103	Overall significant overestimation from cumulative effect of 8XX curricula. No quarterly trend.
P 3104	Large errors in summary predictions generally due to prediction vice accuracy errors. All curricula except 570 and 817 experience these errors. First quarter predictions look good, underestimates occur in the second quarter, overestimates in the third.
P 3105	Significant prediction problems. Generally offered only third quarter.
3302	Not enough data.
3303	Generally good. 525 provides majority of input.
3401	Not enough data.
3402	Prediction problem, not severe.

```
 Course
 Number  Comments
 ******  **************************************************

  3403   Generally excellent, not enough data.

P 3404   Curricula 365 and 825 provide most of the inputs
         to this course. 365 does an excellent job, 825 a
         poor one.

  3601   Generally good.

  3602   Generally excellent.

  3603   Generally excellent.

  3604   Generally good.

P 3702   Significant prediction problem. Not taught first
         quarter, only predicted. Curriculum 847 provides
         most of the input.

  4601   Generally excellent.

  4602   Generally excellent, needs more data.

  4701   Generally good.


************************************************************
**                  END OF COURSE LISTING                 **
************************************************************
```

	Curriculum Number	Comments
	******	**
P	0	Summary predictions are innaccurate due to course effects. No quarterly trend seen.
p	360	Seldom predicted for OS courses. Minor effect due to small number of inputs.
	365	Generally excellent, possible problem with OS3008
	366	Generally excellent.
	367	Generally good, need more data in OS3101.
P	368	Third quarter prediction problem developing.
	372	Not enough data.
	373	Not enough data.
P	374	Tends to overestimate.
	380	Not enough data.
	440	Not enough data.
	441	Not enough data.
P	460	Never predicted.
	525	Generally excellent.
P	530	Problem in OS3104.
P	531	Significant problems affecting OS2103 and OS3104.
	532	Not enough data.
	535	Not enough data.
	570	Possible errors in third quarter, need more data.

	Curriculum Number	Comments
	******	**
P	590	Problem in OS2102
	591	Looks good, not enough data.
	600	Generally good.
	610	Not enough data.
	611	Not enough data.
	620	Generally good.
	683	Not enough data.
	686	Not enough data.
	687	Not enough data.
	813	Generally excellent.
	814	Generally excellent.
	815	Generally fair.
P	817	Problem with OS3006, good otherwise.
P	819	Various prediction problems.
P	825	Problem with OS3404.
P	827	Overestimates for OS3102 and OS3103.
P	837	Problem in OS3102, prediction error for OS3105.
P	847	Problem with OS3702. Good with OS3103. Third quarter predictions are poor.
P	857	Routine innaccuracies with no apparent pattern.
	999	Insignificant input.

APPENDIX D
CONDENSED COURSE MATRICES

This appendix contains the condensed course matrices for all curricula that take OS courses. Curricula 610, 611, 681, 682, 683, and 684 do not have any OS courses required and are not included. 'xxxxxx' entries in the condensed matrices indicate that the students are onboard but not taking OS or OA designated courses (these entries are not made following the final OS or OA course taken to prevent display clutter). The entries are based on the printed course matrices as published as of August, 1984. These matrices were provided by the individual curricular offices.

TABLE IX

Condensed Course Matrices

CURRIC		QUARTER									SCHEDULED ENTRY QTRS
		1	2	3	4	5	6	7	8	9	
365		OS2103	OS3604 OS3404	OS3008	OS3603	OS4602					f
366		CS2103	OS3604 OS3404	CS3008	OS3603 OA3602	OS4602					f
374		xxxxxx	OS2103	CS3604	xxxxxx	xxxxxx	OS3601	OS3602	OS3603		f,w,sp,su
525		xxxxxx	OS2103	CS3303 CS3604	OS3601	xxxxxx	OS3402	OS4601 OS3602			f,sp
535		xxxxxx	xxxxxx	xxxxxx	OS2103						f
595		xxxxxx	OS2103	OS3604	OS3003	OS3603	OS3403	OS4601			f
367		xxxxxx	OS3101	CS3004							f,sp
620		xxxxxx	OS3101	OS3005							f
686		xxxxxx	OS3101								w,su
687		OS3101									w,su
825	A	xxxxxx	OS3101 OS3404	CS3002							f
	B	xxxxxx	OS3101	xxxxxx	OS3404	OS3002					sp
368		OS3001									f,sp
372		xxxxxx	xxxxxx	CS3104							f,w,sp,su
373		xxxxxx	OS3104								f,w,sp,su

Table 9

Condensed Course Matrices (continued)

	QUARTER									SCHEDULED
CURRIC	1	2	3	4	5	6	7	8	9	ENTRY QTRS
******	******	******	******	******	******	******	******	******	******	**********
440	xxxxxx	xxxxxx	CS3104							f,w,sp,su
441	xxxxxx	xxxxxx	CS3104							f,w,sp,su
530	xxxxxx	OS3104								f,sp
531	xxxxxx	OS3104								f,sp
532	xxxxxx	OS3104								f,sp
570	xxxxxx	xxxxxx	xxxxxx	xxxxxx	xxxxxx	xxxxxx	xxxxxx	xxxxxx	OS3104	f,sp
590 A	xxxxxx	xxxxxx	xxxxxx	xxxxxx	xxxxxx	OS2102	OS3104			f,w,sp,su
B	xxxxxx	xxxxxx	xxxxxx	OS2102	xxxxxx	xxxxxx	OS3104			f,w,sp,su
591	xxxxxx	OS2102								f,w,sp,su
600	xxxxxx	xxxxxx	xxxxxx	OS2102						f,w,sp,su
813	xxxxxx	OS3102	CS3103	OS3006						su

Table 9

Condensed Course Matrices (continued)

	QUARTER									SCHEDULED
CURRIC	1	2	3	4	5	6	7	8	9	ENTRY QTRS
******	******	******	******	******	******	******	******	******	******	**********
814	xxxxxx	OS3102	CS3103	OS3006						su
815	xxxxxx	OS3102	CS3103	OS3006						w,su
817	xxxxxx	OS3102	CS3103	OS3006						w,su
(USA)	xxxx	OS3102	OS3103	OS3604 OA2651 OA3201	OA4654	OA4655 -OS3303 -OS3304	or			w,su
819	xxxxxx	OS3102	CS3103 CA3501	OS3006	OA4501					w,su
827	xxxxxx	OS3102	CS3103	OS3006						w,su
837	xxxxxx	OS3102	CS3103	OS3006						w,su
847	xxxxxx	OS3102	CS3103	OS3006 OS3702						w,su
857	xxxxxx	OS3102	CS3103	OS3006						w,su

APPENDIX E

ENTRY / TRACK COURSE DESIGNATIONS

This appendix contains a complete listing of the OS courses divided into 'Entry' and 'Follow-on' catagories. An 'Entry' course is the first OS course that is taken in a given course matrix (for a given curriculum); the first time a curriculum can be 'seen' by the Operations Analysis Department. 'Follow-on' courses are those courses that are taken in quarters following the quarter the 'Entry' course is taken. All curricula take one of the six 'Entry' courses prior to any further OS course work with the exceptions of curricula 610, 611, 681, 682, 683 and 684 which have no required OS courses.

ENTRY COURSE	CURRICULUM	TRACK (Follow-on) COURSES, ALL OS
******	**********	******************************
OS2103	365	3604,3404,3008,3603,4602
	366	3604,3404,3008,3603,4602
	374	3601,3602,3603,3604
	525	3303,3604,3601,3401,4601,3602
	535	none
	595	3604,3003,3603,3403,4601
CS3101	367	3004
	620	3005
	686	none

ENTRY COURSE	CURRICULUM	TRACK (Follow-on) COURSES, ALL OS
******	**********	******************************
CS3101	687	none
	825	3404,3002
OS3001	368	none
OS3104	372	none
	373	none
	440	none
	441	none
	530	none
	531	none
	532	none
	570	none
OS2102	590	3104
	591	none
	600	none
OS3102	813	3103,3006
	814	3103,3006
	815	3103,3006
	817	3103,3006,3604

ENTRY COURSE	CURRICULUM	TRACK (Follow-on) COURSES, ALL OS
******	**********	********************************
OS3102	819	3103,3006
	827	3103,3006
	837	3103,3006
	847	3103,3006,3702
	857	3103,3006

LIST OF REFERENCES

1. Chief of Naval Operations UNCLASSIFIED Letter 1500: 114D/368857 to Commander, Naval Military Personnel Command, Subject: *Fiscal Year 1985 Officer Graduate and Undergraduate Education Quota Plan* 4 June 1984 (NPS copy held by Code 013)

2. *Naval Postgraduate School Catalog, 1983-1984*

3. *NAVPGSCOLINST 1520.19 CH-1*, 21 October 1983

4. McNeil, Donald R., *Interactive Data Analysis*, Wiley, 1977

5. Conover, W. J., *Practical Nonparametric Statistics, Second Edition*, Wiley, 1980

BIBLIOGRAPHY

Duncan, A. J., *Quality Control and Industrial Statistics*, R. D. Irwin, 1974.

Mosteller, F., *Data Analysis and Regression*, Addison-Wesley, 1977

National Bureau of Standards, *Computor Model Documentation Guide*, U.S. Department of Commerce, 1981

Olkin, I., *Probability Models and Applications*, Macmillan, 1980.

Ramsey, J. B., *APL-STAT*, Wadsworth, 1981.

INITIAL DISTRIBUTION LIST

	No. Copies
1. Defense Technical Information Center Cameron Station Alexandria, Va. 22314	2
2. Library, Code 0142 Naval Postgraduate School Monterey, Ca. 93943	2
3. Department Chairman, Code 55 Naval Postgraduate School Monterey, Ca 93943	1
4. Dr. Robert R. Read, Code 55Re Department of Operations Research Naval Postgraduate School Monterey, Ca 93943	1
5. Lcdr Armando R. Solis, Code 55Zb Department of Operations Research Naval Postgraduate School Monterey, Ca 93943	1
6. Lt. Craig C. Madsen, USN Class 86, Department Head Course Surface Warfare Officers School Command Newport, R.I. 02841	4
7. Mr. J. M. Barron, Code 013 Naval Postgraduate School Monterey, Ca. 93943	1
7. Mr. M. Troian, Code 0301 Naval Postgraduate School Monterey, Ca. 93943	1

13537 5

210390

Thesis
M274864 Madsen
c.1 A preliminary analysis of course loading prediction errors.
28 MAY 87 33427

210390

Thesis
M274864 Madsen
c.1 A preliminary analysis of course loading prediction errors.

thesM274864
A preliminary analysis of course loading
3 2768 001 01155 4
DUDLEY KNOX LIBRARY